



---

RESULT 8  
ID Q6NN08 PRELIMINARY; PRT; 190 AA.  
AC DT 05-JUL-2004 (TREMBREL. 27, Created)  
DT 05-JUL-2004 (TREMBREL. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBREL. 27, Last annotation update)  
DE At3g12092  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;  
OC Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; rosids;  
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kim C.J., Chen H., Cheuk R., Shimn P., Carninci P., Hayashizaki Y.,  
RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
RA Seki M., Shinozaki K., Ecker J.R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: BT012169; AAS76264.1; -.  
DR IntezPro: IPR005162; Retrotans\_gag.  
DR Pfam: PF01732; Retrotans\_Gag; 1.  
DR SEQENCE 190 AA; 21466 MW; 821E390B7C17A67B CRC64;  
SQ Query Match 33.3%; Score 56; DB 2; Length 190;  
Best Local Similarity 37.0%; Pred. No. 9;  
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 LDPVKDVLISALRMLWAADDFEDL 27  
Db 157 VEGINDVAAALSALKKALWYKSEFRKEL 183

RESULT 15  
ID AAS76264 PRELIMINARY; PRT; 190 AA.  
AC AAS76264;  
DT 29-MAR-2004 (TREMBREL. 27, Created)  
DT 29-MAR-2004 (TREMBREL. 27, Last sequence update)  
DT 29-MAR-2004 (TREMBREL. 27, Last annotation update)  
DE At3g12092  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;  
OC Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; rosids;  
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kim C.J., Chen H., Cheuk R., Shimn P., Carninci P., Hayashizaki Y.,  
RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
RA Seki M., Shinozaki K., Ecker J.R.;  
RT "Arabidopsis ORF clones";  
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: BT012169; AAS76264.1; -.  
DR SEQENCE 190 AA; 21466 MW; 821E390B7C17A67B CRC64;  
SQ Query Match 33.3%; Score 56; DB 2; Length 190;  
Best Local Similarity 37.0%; Pred. No. 9;  
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 LDPVKDVLISALRMLWAADDFEDL 27  
Db 157 VEGINDVAAALSALKKALWYKSEFRKEL 183



Sequence 2, Application US/08852933  
 Patent No. 5346725  
 GENERAL INFORMATION:  
 APPLICANT: Yokoyama, Shiro  
 TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street, suite 510  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/852,933  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/224,983  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Giulio A. DeConti, Jr.  
 REGISTRATION NUMBER: 31,503  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Giulio A. DeConti, Jr.  
 REGISTRATION NUMBER: 31,503  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7401  
 FILING DATE:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 83 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-852-945-2

Query Match 90.5%; Score 134; DB 2; Length 83;  
 Best Local Similarity 93.5%; Pred. No. 2.6e-13;  
 Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANLILMVPILLAMFLMLTERKILGYIQR 31  
 Db 2 ANLILMVPILLAMFLMLTERKILGYIQR 32

---

RESULT 3  
 US-08-852-945-2  
 Sequence 2, Application US/08852945  
 Patent No. 558674  
 GENERAL INFORMATION:  
 APPLICANT: Yokoyama, Shiro  
 TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street, suite 510  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/853,021  
 FILING DATE: 08-MAY-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/224,983  
 FILING DATE: 08-APR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Giulio A. DeConti, Jr.  
 REGISTRATION NUMBER: 31,503  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 FILING DATE:  
 INFORMATION FOR SEQ ID NO:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 83 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-853-021-2

Query Match 90.5%; Score 134; DB 2; Length 83;  
 Best Local Similarity 93.5%; Pred. No. 2.6e-13;  
 Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANLLLMPILLAMAFMLTTERKILGYQFR 31  
 2 ANLLLIVPILLAMAFMLTTERKILGYQLR 32  
 Db

RESULT 5  
 US-08-852-865-2  
 Sequence 2, Application US/08852865  
 GENERAL INFORMATION:  
 APPLICANT: Yokoyama, Shiro  
 TITLE OF INVENTION: Cisplatin Resistance Gene and Uses Therefor  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street, suite 510  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/852, 865  
 FILING DATE:

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/224, 983  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Giulio A. DeConti, Jr.  
 REGISTRATION NUMBER: 31, 503  
 REFERENCE/DOCKET NUMBER: BBI-010  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 83 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-852-865-2  
 Query Match Score 131; DB 3; Length 83;  
 Best Local Similarity 93.5%; Pred. No. 2.6e-13;  
 Matches 29; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 ANLLLMPILLAMAFMLTTERKILGYQFR 31  
 Db 2 ANLLLIVPILLAMAFMLTTERKILGYQLR 32

RESULT 6  
 US-03-097-889-13  
 Sequence 13, Application US/09097889  
 Patent No. 6218117  
 GENERAL INFORMATION:  
 APPLICANT: Herrnstadt, Corrina  
 APPLICANT: Ghosh, Soumitra S.  
 APPLICANT: Davis, Robert E.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AGENTS THAT QUANTITATIVELY ALTER DETECTABLE MITOCHONDRIAL DNA RATIOS  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington

COUNTRY: USA  
 ZIP: 98104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/097, 889  
 FILING DATE: 15-JUN-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Roseman Ph.D., Stephen J.  
 REGISTRATION NUMBER: 43, 058  
 REFERENCE/DOCKET NUMBER: 660088-417  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-6900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 318 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-09-097-889-13  
 Query Match Score 131; DB 3;  
 Best Local Similarity 90.3%; Pred. No. 3.4e-12;  
 Matches 28; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ANLLLMPILLAMAFMLTTERKILGYQFR 31  
 Db 4 ANLLLIVPILLAMAFMLTTERKILGYQLR 34  
 RESULT 7  
 US-09-098-079-13  
 Sequence 13, Application US/09098079  
 Patent No. 6438095  
 GENERAL INFORMATION:  
 APPLICANT: Herrnstadt, Corrina  
 APPLICANT: Ghosh, Soumitra S.  
 APPLICANT: Cleverger, William  
 APPLICANT: Party, Boin P.  
 APPLICANT: Davis, Robert E.  
 TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/098, 079  
 FILING DATE: 15-JUN-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Roseman Ph.D., Stephen J.  
 REGISTRATION NUMBER: 43, 058  
 REFERENCE/DOCKET NUMBER: 660088-416  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-6900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:

RESULT 10  
US-09-248-796A-14320  
; Sequence 14320, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196-132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIORITY APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIORITY APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 14320  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-14320

Query Match 58.5%; Score 131; DB 4; Length 318;  
Best Local Similarity 90.3%; Pred. No. 3.4e-12;  
Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANLLILVPLILAVAFMLTTERKILGYIOPR 31  
Db 4 ANLLILVPLILAVAFMLTTERKILGYIOPR 34

RESULT 8  
US-09-270-767-41860  
; Sequence 41860, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7322-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 41860  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-41860

Query Match 58.8%; Score 87; DB 4; Length 316;  
Best Local Similarity 60.7%; Pred. No. 1.8e-05;  
Matches 17; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 4 LLLMVPILAVAFMLTTERKILGYIOPR 31  
Db 15 LLLICIVLVSAAFLTLERKILGYIOPR 42

RESULT 9  
US-09-270-767-44872  
; Sequence 44872, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7322-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 44872  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-44872

Query Match 56.1%; Score 83; DB 4; Length 286;  
Best Local Similarity 57.1%; Pred. No. 6.7e-05;  
Matches 16; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 4 LLLMVPILAVAFMLTTERKILGYIOPR 31  
Db 22 LLLICIVLVSAAFLTLERKILGYIOPR 49

RESULT 11  
US-08-928-692-20  
; Sequence 20, Application US/08928692  
; Patent No. 5988727  
; GENERAL INFORMATION:  
; APPLICANT: Brody, Howard S.  
; APPLICATION: Yaver, Deborah S.  
; APPLICANT: Lanska, Michael  
; APPLICANT: Hansen, Kim  
; TITLE OF INVENTION: Methods for Modifying the Production of a Polypeptide  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NC, 5958727, No. 5958727disk of No. 5958727 America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disquette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,692  
; FILING DATE: 12-SEPT-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4944-200-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 20:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5958727e

US-08-928-692-20  
 Query Match 42.6%; Score 63; DB 2; Length 371;  
 Best Local Similarity 41.4%; Pred. No. 0.1;  
 Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
 SEQ ID NO: 41238 LENGTH: 189  
 PRT; TYPE: PRT  
 ORGANISM: Drosophila melanogaster  
 FEATURE:  
 OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-41238

RESULT 12  
 US-09-339-972-20  
 Sequence 20, Application US/09339972  
 ; Patent No. 6323002  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brody, Howard S.  
 ; APPLICANT: Yaver, Deborah S.  
 ; APPLICANT: Lamza, Michael  
 ; APPLICANT: Hansen, Kim  
 ; TITLE OF INVENTION: Methods for Modifying the Production of  
 ; NUMBER OF SEQUENCES: 80  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: No. 63230020 No. 6323002disk of No. 6323002th America, Inc.  
 ; STREET: 405 Lexington Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10174  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATE:  
 ; APPLICATION NUMBER: US/09/339, 972  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/928, 692  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lambiris, Elias J.  
 ; REGISTRATION NUMBER: 33,728  
 ; REFERENCE/DOCKET NUMBER: 4944.200-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-878-9655  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 371 amino acids  
 ; TYPE: amino acid  
 ; STRANGENESS: Single  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE: No. 6323002e  
 US-09-339-972-20

Query Match 42.6%; Score 63; DB 3; Length 371;  
 Best Local Similarity 41.4%; Pred. No. 0.1;  
 Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
 SEQ ID NO: 41238 LENGTH: 189  
 PRT; TYPE: PRT  
 ORGANISM: Drosophila melanogaster  
 FEATURE:  
 OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-41238

RESULT 14  
 US-09-270-767-56454  
 Sequence 56454, Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; CURRENT FILING DATE: 1999-03-17  
 ; NUMBER OF SEQ ID NOS: 62517  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 41238 LENGTH: 189  
 ; TYPE: PRT  
 ; OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-56454

RESULT 15  
 US-09-363-110-4394  
 Sequence 4394, Application US/09583110  
 ; Patent No. 6639703  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al.  
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
 ; FILE REFERENCE: PATH00-07A  
 ; CURRENT APPLICATION NUMBER: US/09/583,110  
 ; CURRENT FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/107,433  
 ; PRIOR FILING DATE: 1998-05-30  
 ; PRIOR APPLICATION NUMBER: US 60/085,131  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: US 60/051,553  
 ; PRIOR FILING DATE: 1997-07-02  
 ; NUMBER OF SEQ ID NOS: 5322 LENGTH: 850  
 ; TYPE: PRT

RESULT 13  
 US-09-270-767-41238  
 Sequence 41238, Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.

? ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4394

Query Match 35.1%; Score 52; DB 4; Length 850;  
Best Local Similarity 33.3%; Pred. No. 13;  
Matches 8; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
Qy 4 LILMVPLIAMAFLMLTERKILGY 27  
Db 161 VFILEPLIITGLHLITEKILLY 184

Search completed: November 10, 2004, 13:44:01  
Job time : 11.7081 secs

Copyright Gencore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model.

Perfect score: 148

Sequence: 1 ANLILMVPILLAMAFMLTERKILGYIQLR 31

Run on: November 10, 2004, 16:36:12 ; Search time 33:30:22 Seconds  
(without alignments)  
328.807 Million cell updates/sec

Title: US-10-092-750-26

Scoring table: BLOSUM62

GAPOP 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6\_ptodata/1/pubbaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6\_ptodata/1/pubbaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6\_ptodata/1/pubbaa/US07\_NEWT\_PUB.pep.\*

4: /cgn2\_6\_ptodata/1/pubbaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6\_ptodata/1/pubbaa/US07\_NEWT\_PUB.pep.\*

6: /cgn2\_6\_ptodata/1/pubbaa/PCITS\_PUBCOMB.pep.\*

7: /cgn2\_6\_ptodata/1/pubbaa/US08\_NEWT\_PUB.pep.\*

8: /cgn2\_6\_ptodata/1/pubbaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6\_ptodata/1/pubbaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6\_ptodata/1/pubbaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6\_ptodata/1/pubbaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6\_ptodata/1/pubbaa/US09\_NNEW\_PUB.pep.\*

13: /cgn2\_6\_ptodata/1/pubbaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6\_ptodata/1/pubbaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6\_ptodata/1/pubbaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6\_ptodata/1/pubbaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6\_ptodata/1/pubbaa/US10\_NNEW\_PUB.pep.\*

18: /cgn2\_6\_ptodata/1/pubbaa/US11\_NNEW\_PUB.pep.\*

19: /cgn2\_6\_ptodata/1/pubbaa/US60\_NNEW\_PUB.pep.\*

20: /cgn2\_6\_ptodata/1/pubbaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	148	100.0	31	14 US-10-092-750-26
2	81.1	88.5	318	9 US-09-098-079-13
3	131	88.5	318	15 US-10-428-487-34
4	131	88.5	318	16 US-10-408-765A-189
5	130	87.8	101	15 US-10-264-049-3171
6	130	87.8	261	17 US-10-425-115-310677
7	120	85.8	318	16 US-10-408-765A-1490
8	120	81.1	79	15 US-10-264-049-3416
9	111.5	75.3	110	15 US-10-424-599-26150
10	65	43.9	355	17 US-10-282-122A-47875
11	64	43.2	46	17 US-10-425-115-332406
12	63	42.6	371	17 US-10-000-845-20
13	60	40.5	134	17 US-10-425-115-202752

## ALIGMENTS

Qy	Db	Length: 31;	Query Match	Best Local Matches	Similarity	Score 148;	DB 14;	Length 31;
				1 ANLLLNVPILLAMAFMLTERKILGYIQLR 31	100.0%; Mismatches 0;	100.0%; Pred. No. 3.3e-14;		
				1 ANLLLNVPILLAMAFMLTERKILGYIQLR 31	100.0%; Mismatches 0;	100.0%; Pred. No. 3.3e-14;		

RESULT 1  
US-10-092-750-26  
; Sequence 26, Application US-10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; APPLICANT: Wright, Martin C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; FILE REFERENCE: 5003650002  
; CURRENT APPLICATION NUMBER: US-10/092,750  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US-60/274,526  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 26  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Homo sapiens

APPLICANT: Herrnstadt, Corrina  
 APPLICANT: Ghosh, Soumitra S.  
 APPLICANT: Clevenger, William  
 APPLICANT: Fahy, Roin F.  
 APPLICANT: Davis, Robert E.  
 TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF NUMBER OF SEQUENCES: 26  
 ADDRESS: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATE: 09-JUN-1998  
 APPLICATION NUMBER: US/09/098,079  
 FILING DATE: 15-JUN-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Roseman Ph.D., Stephen J.  
 REGISTRATION NUMBER: 43 058  
 REFERENCE/DOCKET NUMBER: 660088.416  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 318 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-09-098-079-13

Query Match Score 131; DB 9; Length 318;  
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;  
 Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANLLILVPILLAYAFLMLTERKILGYIOPR 31  
 Db 4 ANLLILVPILLAYAFLMLTERKILGYIOPR 34

RESULT 3  
 US-10-428-487-34  
 / Sequence 34, Application US/10428487  
 / Publication No. US20040006780A1  
 / GENERAL INFORMATION:  
 / APPLICANT: RASTELLI, LUCA K.  
 / APPLICANT: GERBER, HANS-PETER  
 / TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM  
 / FILE REFERENCE: 0980080-0103  
 / CURRENT APPLICATION NUMBER: US/10/428,487  
 / CURRENT FILING DATE: 2003-05-02  
 / PRIOR APPLICATION NUMBER: 09/815,153  
 / PRIOR FILING DATE: 2001-03-21  
 / PRIOR APPLICATION NUMBER: 60/191,201  
 / PRIOR FILING DATE: 2000-01-22  
 / NUMBER OF SEQ ID NOS: 84  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO: 34  
 / LENGTH: 318  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / US-10-428-487-34

Query Match Score 131; DB 15; Length 318;  
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;

Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 ANLLILVPILLAYAFLMLTERKILGYIOPR 31  
 Db 4 ANLLILVPILLAYAFLMLTERKILGYIOPR 34  
 RESULT 4  
 US-10-408-765A-189  
 / Sequence 189, Application US/10408765A  
 / Publication No. US20040101874A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Ghosh, Soumitra S.  
 / APPLICANT: Fahy, Roin D.  
 / APPLICANT: Zhang, Bing  
 / APPLICANT: Gibson, Bradford W.  
 / APPLICANT: Taylor, Steven W.  
 / APPLICANT: Glenn, Gary M.  
 / APPLICANT: Warnock, Dale E.  
 / TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 / TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 / FILE REFERENCE: 660088.465  
 / CURRENT APPLICATION NUMBER: US/10/408,765A  
 / CURRENT FILING DATE: 2003-04-04  
 / NUMBER OF SEQ ID NOS: 3077  
 / SOFTWARE: FastSEQ for Windows Version 4.0  
 / SEQ ID NO: 189  
 / LENGTH: 318  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / US-10-408-765A-189  
 Query Match Score 131; DB 16; Length 318;  
 Best Local Similarity 90.3%; Pred. No. 1.2e-10;  
 Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 ANLLILVPILLAYAFLMLTERKILGYIOPR 31  
 Db 4 ANLLILVPILLAYAFLMLTERKILGYIOPR 34  
 RESULT 5  
 US-10-264-049-171  
 / Sequence 3171, Application US/10264049  
 / Publication No. US2004005579A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Birse et al.  
 / TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 / FILE REFERENCE: PA133P1  
 / CURRENT APPLICATION NUMBER: US/10/264,049  
 / CURRENT FILING DATE: 2002-10-04  
 / PRIOR APPLICATION NUMBER: PCT/US01/18569  
 / PRIOR FILING DATE: 2001-05-07  
 / PRIOR APPLICATION NUMBER: US 60/209,467  
 / PRIOR FILING DATE: 2000-06-07  
 / NUMBER OF SEQ ID NOS: 4360  
 / SOFTWARE: PatentIn Ver. 3.1  
 / SEQ ID NO: 3171  
 / LENGTH: 101  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / FEATURE: MISC\_FEATURE  
 / LOCATION: (3)  
 / NAME/KEY: MISC\_FEATURE  
 / OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 / FEATURE:  
 / NAME/KEY: MISC\_FEATURE  
 / LOCATION: (1)  
 / OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 / FEATURE:  
 / NAME/KEY: MISC\_FEATURE  
 / LOCATION: (4)  
 / OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 /

```

FEATURE;
NAME/KEY: MISC_FEATURE
LOCATION: (5)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE;
NAME/KEY: MISC_FEATURE
LOCATION: (7)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
JS-10-264-049-3171

Query Match Score 130; DB 15; Length 101;
Best Local Similarity 93.3%; Pred. No. 4_8e-11;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DQY 2 NLLLIVPILIANAFLMLTERKILGYIQR 31
DOb 21 NLLLIVPILIANAFLMLTERKILGYIQR 50

RESULT 6
JS-10-425-115-310657
Sequence 310657, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 363326
SEQ ID NO 310657
LENGTH: 261
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) .. (261)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46377C.1.pep
US-10-425-115-310657

Query Match Score 130; DB 17; Length 261;
Best Local Similarity 93.3%; Pred. No. 1_3e-10;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DQY 2 NLLLIVPILIANAFLMLTERKILGYIQR 31
DOb 1 NLLLIVPILIANAFLMLTERKILGYIQR 30

RESULT 7
US-10-408-765A-1490
Sequence 1490, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3027

```

Query Match Similarity 75.3%; Score 111.5; DB 15; Length 110;  
 Best Local Similarity 84.4%; Pred. No. 2.5e-08;  
 Matches 27; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; Title of Invention: Plants  
 ; FILE REFERENCE: 38-211532221.B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO: 332406  
 ; LENGTH: 46  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)-(46)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_6626C.1.pep  
 ; US-10-425-115-332406

Query Match Similarity 43.2%; Score 64; DB 17; Length 46;  
 Best Local Similarity 82.4%; Pred. No. 0.069;  
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 AMAFLMLTTERKILGTYQ 29  
 Db 1 AMAFLMLTTERKILGTYQ 17

RESULT 12  
 US-10-000-845-20  
 ; Sequence 20, Application US/10000845  
 ; Publication No. US20040197854A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brody, Howard  
 ; Yaver, Deborah S.  
 ; Lamsa, Michael  
 ; Hansen, Kim  
 ; TITLE OF INVENTION: Methods for Modifying the Production of  
 ; NUMBER OF SEQUENCES: 80  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Novo Nordisk of North America, Inc.  
 ; STREET: 405 Lexington Avenue  
 ; CITY: New York  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10174  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ FOR Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/000,845  
 ; FILING DATE: 24-OCT-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/000845  
 ; FILING DATE: 12-SEPT-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lambiris, Elias J  
 ; REGISTRATION NUMBER: 33,728  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-967-0123  
 ; TELEX/FAX: 212-878-9655  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 371 amino acids

Query Match Similarity 43.9%; Score 65; DB 15; Length 355;  
 Best Local Similarity 41.4%; Pred. No. 0.46;  
 Matches 12; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

GENERAL INFORMATION:  
 ; APPLICANT: Burkholderia cepacia  
 ; ORGANISM: Burkholderia cepacia  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO: 47875  
 ; LENGTH: 355  
 ; TYPE: PRT  
 ; ORGANISM: Burkholderia cepacia  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-211532221.B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.

RESULT 11  
 US-10-425-115-332406  
 ; Sequence 332406, Application US/10425115  
 ; Publication No. US20040214272A1

Query Match Similarity 43.9%; Score 65; DB 15; Length 355;  
 Best Local Similarity 41.4%; Pred. No. 0.46;  
 Matches 12; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

GENERAL INFORMATION:  
 ; APPLICANT: Ilvsvvllcvayllwerklgmhvrv  
 ; SEQ ID NO: 28  
 ; LENGTH: 355

TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: Linear  
 MOLECULE TYPE: None  
 SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
 US-10-000-845-20

Query Match Score 63; DB 17; Length 371;  
 Best Local Similarity 42.6%; Pred. No. 0.93;  
 Matches 12; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 3 LLLIMVPILJAMAFLMLTMRKILGYIOPR 31  
 Db 13 VLLVLPSSLLAVTVVAAERKTMASMQRR 41

RESULT 13  
 US-10-425-115-202752  
 Sequence 202752, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 202752  
 LENGTH: 134  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1) .. (134)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE: OTHER INFORMATION: Clone ID: MRT4577\_11649C.1.pep  
 US-10-425-115-202752

Query Match Score 60; DB 17; Length 134;  
 Best Local Similarity 37.9%; Pred. No. 0.84;  
 Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 LLLIMVPILJAMAFLMLTMRKILGYIOPR 31  
 Db 13 VLLVLPSSLLAVTVVAAERKTMASMQRR 41

RESULT 14  
 US-10-424-599-161174  
 Sequence 161174, Application US/10424599  
 Publication No. US2004031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO 161174  
 LENGTH: 138  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 NAME/KEY: unsure

LOCATION: (1) .. (138)  
 OTHER INFORMATION: unsure at all Xaa locations  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_116558C.1.pep  
 US-10-424-599-161174

Query Match Score 60; DB 15; Length 138;  
 Best Local Similarity 37.9%; Pred. No. 0.86;  
 Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 3 LLLIMVPILJAMAFLMLTMRKILGYIOPR 31  
 Db 13 VLLVLPSSLLAVTVVAAERKTMASMQRR 41

RESULT 15  
 US-10-282-122A-45946  
 Sequence 45946, Application US/1028212A1  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cherry  
 APPLICANT: Hasebeck, Robert  
 APPLICANT: Olsen, Kari  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: EIJTRA\_034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/19,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 45946  
 LENGTH: 261

TYPE: PRT  
 ORGANISM: Bacillus anthracis  
 FEATURE:  
 NAME/KEY: (66)-(66)  
 LOCATION: (66)-(66)  
 FEATURE:  
 NAME/KEY: MISC\_FEATURE  
 LOCATION: (62)-(82)  
 OTHER INFORMATION: X=any amino acid  
 FEATURE:  
 NAME/KEY: MISC\_FEATURE

? LOCATION: (197)..(197)  
? OTHER INFORMATION: X=any amino acid  
US-10-282-122A-45946

Query Match 37.2%; Score 55; DB 15; Length 261;  
Best Local Similarity 54.5%; Pred. No. 9.1;  
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 5 LIMVPILIAAFMLTERKIG 26  
| | : | | | : |||  
Db 112 LLMATYYVANGFLYINETRIG 133

Search completed: November 11, 2004, 01:28:08  
Job time : 34.3522 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using SW mode.

Run on: November 10, 2004, 12:32:37 ; Search time 6.90846 Seconds  
(without alignments)

Title: US-10-092-750-40

Perfect score: 106

Sequence: 1 SSQARIHQWLHFLFDDFTST 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing First 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/1aa/'5A COMB pep'\*  
2: /cgn2\_6/ptodata/1/1aa/'5B COMB pep'\*  
3: /cgn2\_6/ptodata/1/1aa/'6A COMB pep'\*  
4: /cgn2\_6/ptodata/1/1aa/'6B COMB pep'\*  
5: /cgn2\_6/ptodata/1/1aa/'PETUS COMB pep'\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	50	47.2	156	4	US-09-248-796A-18650	Sequence 18650, A
2	46	43.4	105	4	US-09-328-552-4817	Sequence 4817, AP
3	45	42.5	56	4	US-09-621-576-5966	Sequence 5966, AP
4	43.5	41.0	229	4	US-09-252-591A-18086	Sequence 18086, A
5	43	40.6	193	4	US-09-448-796A-18249	Sequence 18249, A
6	43	40.6	301	4	US-09-252-591A-24016	Sequence 24016, A
7	43	40.6	402	4	US-09-270-767-42277	Sequence 42272, A
8	43	40.6	534	4	US-09-293-519-8	Sequence 8, Appl1
9	42	39.6	594	4	US-09-107-532A-5929	Sequence 692, AP
10	41	38.7	63	4	US-09-248-796A-23739	Sequence 23739, A
11	41	38.7	93	4	US-09-448-796A-33391	Sequence 23391, A
12	41	38.7	115	4	US-09-513-939C-4944	Sequence 4944, AP
13	41	38.7	216	3	US-08-464-052-6	Sequence 6, Appl1
14	41	38.7	216	3	US-08-461-002-2	Sequence 6, Appl1
15	41	38.7	216	3	US-08-589-411-6	Sequence 2, Appl1
16	41	38.7	216	4	US-08-192-210-6	Sequence 2, Appl1
17	41	38.7	240	4	US-09-270-767-33620	Sequence 33620, A
18	41	38.7	472	4	US-09-134-000C-5541	Sequence 5541, AP
19	41	38.7	511	3	US-08-464-052-2	Sequence 2, Appl1
20	41	38.7	511	3	US-08-461-002-2	Sequence 2, Appl1
21	41	38.7	511	3	US-08-619-411-2	Sequence 2, Appl1
22	41	38.7	511	4	US-08-392-210-2	Sequence 2, Appl1
23	41	38.7	511	5	PCT-US94-09863-2	Sequence 2, Appl1
24	41	38.7	999	4	US-09-147-371-2	Sequence 2, Appl1
25	40.5	38.2	330	1	US-08-468-833-6	Sequence 6, Appl1
26	40.5	38.2	330	1	US-08-468-835-6	Sequence 6, Appl1
27	40.5	38.2	330	1	US-08-310-357-6	Sequence 6, Appl1

#### RESULT 1

US-09-248-796A-18650

```
; Sequence 18650, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;   APPLICANT: Keith Weinstock et al
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;   TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 107196-132
;   CURRENT APPLICATION NUMBER: US/09/248,796A
;   CURRENT FILING DATE: 1999-02-12
;   PRIOR APPLICATION NUMBER: US 60/074,725
;   PRIOR FILING DATE: 1998-02-13
;   PRIORITY APPLICATION NUMBER: US 60/096,409
;   PRIORITY FILING DATE: 1998-08-13
;   NUMBER OF SEQ ID NOS: 28208
;   SEQ ID NO 18650
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-18650
```

Query Match 47.2%; Score 50; DB 4; Length 156;

```
; Best Local Similarity 52.6%; Pred. No. 1-3;
; Matches 10; Conservative 3; Mismatches 6; Indels 0; Caps 0;
Qy 1 SSQALRTHQWLHFLFSDFTST 19
Db 71 SFGVQTHTSLLHFKNFTS 89
```

#### RESULT 2

US-09-328-352-4877

```
; Sequence 4877, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;   APPLICANT: Gary L. Breton et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;   FILE REFERENCE: GTC99-01PA
;   CURRENT APPLICATION NUMBER: US/09/328,352
;   CURRENT FILING DATE: 1999-06-04
;   NUMBER OF SEQ ID NOS: 8252
;   SEQ ID NO 4877
;   LENGTH: 1005
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-4877
```

```
; Query Match 43.4%; Score 46; DB 4; Length 1005;
```

Best Local Similarity 53.8%; Pred. No. 42;  
 Matches 7; Conservative 2; Mismatches 4;  
 Indels 0; Gaps 0;

Qy 2 SQARLHQWLIF 14  
 Db 691 SQGFHVHQWLAMF 703

RESULT 3  
 US-09-621-976-5966  
 ; Sequence 5966, Application US/09621976  
 ; GENERAL INFORMATION:  
 ; PATENT NO. 6339063  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; INVENTION: ETS and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET-0544R2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5966  
 ; LENGTH: 56  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-621-976-5966

Query Match 42.5%; Score 45; DB 4; Length 56;

Best Local Similarity 66.7%; Pred. No. 2.7;  
 Matches 8; Conservative 1; Mismatches 3;  
 Indels 0; Gaps 0;

Qy 2 SQARLHQWLH 13  
 Db 4 SQATRTHSWYL 15

RESULT 4  
 US-09-252-991A-18086  
 ; Sequence 18086, Application US/09252991A  
 ; PATENT NO. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196-136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US/09/252,991A  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US/094,190  
 NUMBER OF SEQ ID NOS: 33142

Query Match 40.6%; Score 43; DB 4; Length 301;  
 Best Local Similarity 77.8%; Pred. No. 35;  
 Matches 7; Conservative 1; Mismatches 1;  
 Indels 0; Gaps 0;

Qy 3 QALRTHOWL 11  
 Db 76 QLLRLHQWL 84

RESULT 5  
 US-09-252-991A-18086  
 ; Sequence 18086, Application US/09252991A  
 ; PATENT NO. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: File Reference: 7236-094  
 CURRENT APPLICATION NUMBER: US/09/270,767  
 CURRENT FILING DATE: 1999-03-17  
 NUMBER OF SEQ ID NOS: 62217  
 SOFTWARE: PatentIn Ver. 2.0  
 LENGTH: 402  
 TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster

RESULT 6  
 US-09-252-991A-24016  
 ; Sequence 24016, Application US/09252991A  
 ; PATENT NO. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196-136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 24016  
 ; LENGTH: 301  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 18086  
 ; LENGTH: 239  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa

RESULT 7  
 US-09-270-767-42272  
 ; Sequence 42272, Application US/09270767  
 ; PATENT NO. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hamburger et al.  
 ; FILE REFERENCE: File Reference: 7236-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; CURRENT FILING DATE: 1999-03-17  
 ; NUMBER OF SEQ ID NOS: 62217  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; LENGTH: 402  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster

US-09-270-767-42272

Query Match 40.6%; Score 43; DB 4; Length 402;  
 Best Local Similarity 56.2%; Pred. No. 48; Indels 0;  
 Matches 9; Conservative 3; Mismatches 4; Gaps 0;  
 SEQ ID NO: 4 ALRIHOWLHLFSDFTS 19  
 :|||:|||:|||:  
 Db 271 SLLIHL-SKLHFANFTS 286

RESULT 8  
 US-09-293-549-8  
 ; Sequence 8, Application US/09293549  
 ; Patent No. 6440409  
 ; GENERAL INFORMATION:  
 ; APPLICANT: G. Todd Milne  
 ; ATTORNEY/AGENT INFORMATION:  
 ; TITLE OF INVENTION: A METHOD FOR EFFICIENT AND HIGHLY  
 ; SELECTIVE CONTROL OF MICROORGANISMS  
 ; FILE REFERENCE: 50078/008002  
 ; CURRENT FILING DATE: 1999-04-16  
 ; EARLIER FILING NUMBER: 60/082,089  
 ; EARLIER FILING DATE: 1998-04-17  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO: 8  
 ; LENGTH: 534  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 US-09-293-549-8

Query Match 40.6%; Score 43; DB 4; Length 534;  
 Best Local Similarity 50.0%; Pred. No. 65; Indels 0;  
 Matches 10; Conservative 2; Mismatches 8; Gaps 0;  
 SEQ ID NO: 1 SSQALRIHOWLHLFSDFTST 20  
 :|||:|||:|||:  
 Db 482 SASALQNRSNKGGLGSDFT 501

RESULT 9  
 US-09-107-532A-6929  
 ; Sequence 6929, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A. Doucette-Stamm  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02454  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107-532A  
 ; FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: STC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 893-8277  
 TELEFAX: (781) 893-8277  
 INFORMATION FOR SEQ ID NO: 6929:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 594 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (B) 104-11594  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6929:  
 ; Query Match 39.6%; Score 42; DB 4; Length 594;  
 ; Best Local Similarity 46.7%; Pred. No. 1.1e+12;  
 ; Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 ; US-09-107-532A-6929  
 ; RESULT 10  
 ; Sequence 23739, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; CURRENT FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,755  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; PRIOR FILING DATE: 1998-08-13  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO: 23739  
 ; LENGTH: 63  
 ; TYPE: PRT  
 ; ORGANISM: Candida albicans  
 ; US-09-248-796A-23739  
 ; Query Match 38.7%; Score 41; DB 4; Length 63;  
 ; Best Local Similarity 58.3%; Pred. No. 14;  
 ; Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 ; RESULT 11  
 ; Sequence 23391, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; CURRENT FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIORITY FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO: 23391  
 LENGTH: 93  
 TYPE: PRT  
 ORGANISM: *Candida albicans*  
 US-09-248-796A-23391

Query Match 38.7%; Score 41; DB 4; Length 93;  
 Best Local Similarity 53.3%; Pred. No. 21;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 SEQ 5 LRIHQWHLFSDFTS 19  
 :|:||:||:||:  
 Db 73 LNVAQEFVLFSNFTS 87

RESULT 12  
 US-09-513-999C-4944  
 / Sequence 4944, Application US/09513999C  
 / Patent No. 6783961  
 / GENERAL INFORMATION:  
 / APPLICANT: Dumas Milne Edwards, J.B.  
 / APPLICANT: Duclert, A.  
 / APPLICANT: Giordano, J.Y.  
 / TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 / Patent No. 6783961  
 / FILE REFERENCE: 59 US2, REG  
 / CURRENT APPLICATION NUMBER: US/09/513,999C  
 / CURRENT FILING DATE: 2000-02-24  
 / PRIOR APPLICATION NUMBER: US 60/122,487  
 / PRIOR FILING DATE: 1999-02-26  
 / SOFTWARE: Patent\_pm  
 / SEQ ID NO 4944  
 / LENGTH: 115  
 / TYPE: PRT  
 / OTHER INFORMATION: Xaa=Pro or Thr  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: UNSURE  
 / LOCATION: 25  
 / OTHER INFORMATION: Xaa=Pro or Thr  
 US-09-513-999C-4944

Query Match 38.7%; Score 41; DB 4; Length 115;  
 Best Local Similarity 31.6%; Pred. No. 26;  
 Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 SEQ 1 SSQARLHQWHLFSDFTS 19  
 :|:||:||:||:  
 Db 94 ASRIMNYQFQLQYKDTI 112

RESULT 13  
 US-08-464-052-6  
 / Sequence 6, Application US/08464052  
 / Patent No. 6008201  
 / GENERAL INFORMATION:  
 / APPLICANT: Riley M.D., Lee W.  
 / TITLE OF INVENTION: DNA Molecule Encoding for Cellular  
 / NUMBER OF SEQUENCES: 10  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
 / STREET: Clinton Square, P.O. Box 1051  
 / CITY: Rochester  
 / STATE: New York  
 / ZIP: 14603  
 / COUNTRY: U.S.A.

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 APPLICATION NUMBER: US/08/461,002  
 FILING DATE:  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Goldmar, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/186 (D-1485B)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1600  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 216 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 US-08-464-052-6

Query Match 38.7%; Score 41; DB 3; Length 216;  
 Best Local Similarity 47.1%; Pred. No. 51;  
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
 US-08-461-002-6

Query Match 38.7%; Score 41; DB 3; Length 216;  
 Best Local Similarity 47.1%; Pred. No. 51;  
 Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
 US-08-461-002-6

Qy 4 ALRIHOWLHFSDFST 20  
 Db 29 AQRVMDWLHPGDLDT 45

---

RESULT 15  
 US-08-689-411-6  
 Sequence 6, Application US/08689411  
 Patent No. 624881

GENERAL INFORMATION:  
 APPLICANT: Riley, M.D., Lee W.  
 APPLICANT: Chong, Pele  
 TITLE OF INVENTION: DNA MOLECULE FRAGMENTS ENCODING FOR  
 TITLE OF INVENTION: CELLULAR UPTAKE OF MYCOBACTERIUM TUBERCULOSIS AND USES  
 TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
 STREET: Clinton Square, P.O. Box 1051  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/689,411  
 FILING DATE:  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/187

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1600  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 216 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide

US-08-689-411-6

Query Match Score 41; DB 3; Length 216;  
 Best Local Similarity 47.1%; Pred. No. 51;  
 Matches 8; Conservative 1; Nismatches 8; Indels 0; Gaps 0;

Qy 4 ALRIHOWLHFSDFST 20  
 Db 29 AQRVMDWLHPGDLDT 45

Search completed: November 10, 2004, 13:44:10  
 Job time : 6.90846 sec

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw mode!

Run on: November 10, 2004, 16:36:12 ; Search time 21.4953 Seconds  
 (without alignments) 328.807 Million cell. updates/sec

Title: US-10-092-750-40

Perfect score: 106

Sequence: 1 SSQALRTHQWLHLSDFST 20

Scoring table: BLOSUM62

Gapext 0.5

Searched: 1566620 seqs, 35325886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0‡

Maximum Match 100‡

Listing first 45 summaries

Database : Published\_Applications\_AA:<sup>\*</sup>

1: /cgn2\_6\_ptodata/1/pubpaas/US07\_PUBCOMB.pep:\*

2: /cgn2\_6\_ptodata/1/pubpaas/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6\_ptodata/1/pubpaas/US06\_NEV\_PUB.pep:\*

4: /cgn2\_6\_ptodata/1/pubpaas/US06\_PUBCOMB.pep:\*

5: /cgn2\_6\_ptodata/1/pubpaas/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6\_ptodata/1/pubpaas/PCTU\_PUBCOMB.pep:\*

7: /cgn2\_6\_ptodata/1/pubpaas/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6\_ptodata/1/pubpaas/US08\_PUBCOMB.pep:\*

9: /cgn2\_6\_ptodata/1/pubpaas/US09\_PRBCOMB.pep:\*

10: /cgn2\_6\_ptodata/1/pubpaas/US09B\_PRBCOMB.pep:\*

11: /cgn2\_6\_ptodata/1/pubpaas/US09C\_PRBCOMB.pep:\*

12: /cgn2\_6\_ptodata/1/pubpaas/US09\_New\_PUB.pep:\*

13: /cgn2\_6\_ptodata/1/pubpaas/US10A\_PRBCOMB.pep:\*

14: /cgn2\_6\_ptodata/1/pubpaas/US10B\_PRBCOMB.pep:\*

15: /cgn2\_6\_ptodata/1/pubpaas/US10C\_PRBCOMB.pep:\*

16: /cgn2\_6\_ptodata/1/pubpaas/US10D\_PRBCOMB.pep:\*

17: /cgn2\_6\_ptodata/1/pubpaas/US10\_NEW\_PUB.pep:\*

18: /cgn2\_6\_ptodata/1/pubpaas/US11\_NEW\_PUB.pep:\*

19: /cgn2\_6\_ptodata/1/pubpaas/US11\_New\_PUB.pep:\*

20: /cgn2\_6\_ptodata/1/pubpaas/US00\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query	Match	Length	DB ID	Description
1	106	100.0	20	14	US-10-092-750-40	Sequence 40, Appl
2	47	44.3	132	17	US-10-092-115-317501	Sequence 317501, Appl
3	47	44.3	675	14	US-10-302-267-42	Sequence 42, Appl
4	47	44.3	675	15	US-10-412-699B-704	Sequence 704, Appl
5	46	43.4	369	14	US-10-104-047-2357	Sequence 2357, Appl
6	45.5	42.9	175	17	US-10-425-115-341865	Sequence 341865, Appl
7	45	42.5	238	9	US-09-727-855B-9	Sequence 9, Appl
8	45	42.5	335	15	US-10-425-114-38355	Sequence 38355, Appl
9	45	42.5	460	15	US-10-425-114-70131	Sequence 70131, Appl
10	44	41.5	83	17	US-10-425-115-359581	Sequence 359581, Appl
11	44	41.5	126	17	US-10-425-115-345444	Sequence 345444, Appl
12	44	41.5	173	15	US-10-424-599-232233	Sequence 232233, Appl
13	44	41.5	184	15	US-10-424-599-235750	Sequence 235750, Appl

**ALIGNMENTS**

```

RESULT 1
US-10-092-750-40
; Sequence 40, Application US/10092750
; Publication No. US0030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Phillip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Peptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 40
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-40

```

; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO: 311501  
 ; LENGTH: 132  
 ; TYPE: PRT  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_52639C.1.pep  
 us-10-425-115-317501  
 ; Query Match 44.3%; Score 47; DB 17; Length 132;  
 ; Best Local Similarity 87.5%; Pred. No. 17;  
 ; Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 ;  
 QY 8 HOWLHFT 15  
 Db 102 HOWLHFT 109  
 ;  
 RESULT 3  
 us-10-302-267-42  
 ; Sequence 42, Application US10302267  
 ; Publication No. US200229915A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keddie, James  
 ; APPLICANT: Fromm, Michael  
 ; APPLICANT: Heard, Jacqueline  
 ; APPLICANT: Riechmann, Jose Luis  
 ; APPLICANT: Broun, Pierre  
 ; APPLICANT: Pineda, Omaira  
 ; APPLICANT: Reuber, Lynne  
 ; APPLICANT: Yu, Guo-Liang  
 ; APPLICANT: Jiang, Cai-Zhong  
 ; APPLICANT: Samaha, Raymond R.  
 ; APPLICANT: Pilgrim, Marsha L.  
 ; APPLICANT: Creelman, Robert A.  
 ; APPLICANT: Dubell, Arnold N.  
 ; APPLICANT: Ratcliffe, Oliver  
 ; APPLICANT: Kumamoto, Roderick K.  
 ; APPLICANT: Sherman, Bradley K.  
 ; FILE REFERENCE: MBI-0048CIP  
 ; CURRENT APPLICATION NUMBER: US10/412,699B  
 ; CURRENT FILING DATE: 2003-04-10  
 ; PRIOR APPLICATION NUMBER: 09/394,519  
 ; PRIOR FILING DATE: 1999-09-13  
 ; PRIOR APPLICATION NUMBER: 09/489,376  
 ; PRIOR FILING DATE: 2000-01-21  
 ; PRIOR APPLICATION NUMBER: 09/505,720  
 ; PRIOR FILING DATE: 2000-02-17  
 ; PRIOR APPLICATION NUMBER: 09/533,030  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,392  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,029  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,591  
 ; PRIOR FILING DATE: 2001-03-27  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 2011  
 ; SOFTWARE: PatentIn Version 3.2  
 ; SEQ ID NO: 704  
 ; LENGTH: 675  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; OTHER INFORMATION: G1055  
 us-10-412-699B-704  
 ; Query Match 44.3%; Score 47; DB 14; Length 675;  
 ; Best Local Similarity 58.3%; Pred. No. 85;  
 ; Matches 7; Conservative 2; Mismatches 3;  
 ; Indels 0;  
 ; Gaps 0;  
 QY 1 SSQALRHQWHL 12  
 Db 510 SSEGRRLHQWFH 521

; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; OTHER INFORMATION: G1055  
 us-10-302-267-42  
 ; Query Match 44.3%; Score 47; DB 14; Length 675;  
 ; Best Local Similarity 58.3%; Pred. No. 85;  
 ; Matches 7; Conservative 2; Mismatches 3;  
 ; Indels 0;  
 ; Gaps 0;  
 QY 1 SSQALRHQWHL 12  
 Db 510 SSEGRRLHQWFH 521

; RESULT 4  
 us-10-412-699B-704  
 ; Sequence 704, Application US10412699B  
 ; Publication No. US20040045049A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendel Biotechnology, Inc.  
 ; APPLICANT: Zhang, James  
 ; APPLICANT: Fromm, Michael E.  
 ; APPLICANT: Heard, Jacqueline B.  
 ; APPLICANT: Riechmann, Jose Luis  
 ; APPLICANT: Adam, Luc J.  
 ; APPLICANT: Broun, Pierre E.  
 ; APPLICANT: Pineda, Omaira  
 ; APPLICANT: Reuber, Lynne  
 ; APPLICANT: Keddie, James S.  
 ; APPLICANT: Yu, Guo-Liang  
 ; APPLICANT: Jiang, Cai-Zhong  
 ; APPLICANT: Samaha, Raymond R.  
 ; APPLICANT: Pilgrim, Marsha L.  
 ; APPLICANT: Creelman, Robert A.  
 ; APPLICANT: Dubell, Arnold N.  
 ; APPLICANT: Ratcliffe, Oliver  
 ; APPLICANT: Kumamoto, Roderick K.  
 ; APPLICANT: Sherman, Bradley K.  
 ; FILE REFERENCE: MBI-0048CIP  
 ; CURRENT APPLICATION NUMBER: US10/412,699B  
 ; CURRENT FILING DATE: 2003-04-10  
 ; PRIOR APPLICATION NUMBER: 09/394,519  
 ; PRIOR FILING DATE: 1999-09-13  
 ; PRIOR APPLICATION NUMBER: 09/489,376  
 ; PRIOR FILING DATE: 2000-01-21  
 ; PRIOR APPLICATION NUMBER: 09/505,720  
 ; PRIOR FILING DATE: 2000-02-17  
 ; PRIOR APPLICATION NUMBER: 09/533,030  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,392  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,029  
 ; PRIOR FILING DATE: 2000-03-22  
 ; PRIOR APPLICATION NUMBER: 09/533,591  
 ; PRIOR FILING DATE: 2001-03-27  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 2011  
 ; SOFTWARE: PatentIn Version 3.2  
 ; SEQ ID NO: 704  
 ; LENGTH: 675  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; OTHER INFORMATION: G1055  
 us-10-412-699B-704

Query Match      Score 47; DB 15; Length 675;  
 Best Local Similarity 58.3%; Pred. No. 85;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy      1 SSQALRTHMLH 12  
 Db      510 SSEGRGLRHWFH 521

RESULT 5  
 US-10-104-047-2357  
 ; Sequence 2357, Application US/10104047  
 ; Publication No. US20030236392A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: HELIX RESEARCH INSTITUTE  
 ;   TITLE OF INVENTION: No. US20030236392A1el full length cDNA  
 ;   CURRENT APPLICATION NUMBER: US/10/104-047  
 ;   CURRENT FILING DATE: 2002-03-25  
 ;   PRIOR APPLICATION NUMBER:  
 ;   PRIOR FILING DATE:  
 ;   NUMBER OF SEQ ID NOS: 4096  
 ;   SOFTWARE: PatentIn Ver. 2.1  
 ;   SSQ ID NO: 2357  
 ;   TYPE: PRT  
 ;   ORGANISM: Homo sapiens  
 US-10-104-047-2357

Query Match      Score 46; DB 14; Length 369;  
 Best Local Similarity 61.5%; Pred. No. 67;  
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy      3 QALRIHQWLHFS 15  
 Db      17 RALSHRWLWLSFS 29

RESULT 6  
 US-10-425-115-341865  
 ; Sequence 341865, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: La Rosa, Thomas J.  
 ;   APPLICANT: Kovacic, David J.  
 ;   APPLICANT: Zhou, Yihua  
 ;   APPLICANT: Cao, Yongwei  
 ;   TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ;   TITLE OF INVENTION: Plants  
 ;   FILE REFERENCE: 38-21(5322)B  
 ;   CURRENT APPLICATION NUMBER: US/10/425,115  
 ;   CURRENT FILING DATE: 2003-04-28  
 ;   NUMBER OF SEQ ID NOS: 369326  
 ;   SEQ ID NO: 341865  
 ;   TYPE: PRT  
 ;   FEATURE: Zea mays  
 ;   OTHER INFORMATION: Clone ID: MRT4577\_74948C.1.pep  
 US-10-425-115-341865

Query Match      Score 45.5%; DB 17; Length 175;  
 Best Local Similarity 55.0%; Pred. No. 38;  
 Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;  
 Qy      1 SSQA-IRIHQWLHFSDFS 19  
 Db      151 SNQAVMVRGSGWLLFCDFCS 170

RESULT 7  
 US-09-727-855B-9  
 ; Sequence 9, Application US/09727855B  
 ; OTHER INFORMATION: Nucleic Acid Molecules and Other Molecules Associated With

Query Match      Score 45.5%; DB 9; Length 238;  
 Best Local Similarity 80.0%; Pred. No. 62;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy      7 THQWLHFS 16  
 Db      74 VHQLHLFSD 83

RESULT 8  
 US-10-425-114-38355  
 ; Sequence 38355, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: Liu, Jingdong  
 ;   APPLICANT: Zhou, Yihua  
 ;   APPLICANT: Kovacic, David K.  
 ;   APPLICANT: Screen, Steven E  
 ;   APPLICANT: Tabaska, Jack E  
 ;   APPLICANT: Cao, Yongwei  
 ;   TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ;   TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ;   FILE REFERENCE: 38-21(5331)B  
 ;   CURRENT APPLICATION NUMBER: US/10/425,114  
 ;   CURRENT FILING DATE: 2003-04-28  
 ;   NUMBER OF SEQ ID NOS: 73128  
 ;   SEQ ID NO: 38355  
 ;   LENGTH: 335  
 ;   TYPE: PRT  
 ;   ORGANISM: Zea mays  
 ;   FEATURE:  
 ;   OTHER INFORMATION: Clone ID: 700051738\_FLI.pep  
 US-10-425-114-38355

Query Match      Score 45; DB 15; Length 335;  
 Best Local Similarity 57.1%; Pred. No. 87;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Qy      3 QALRIHQWLHFS 16  
 Db      113 QALDHQHVAFPD 126

RESULT 9  
 US-10-425-114-70131  
 ; Sequence 70131, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ;   APPLICANT: Zhou, Yihua  
 ;   APPLICANT: Kovacic, David K.  
 ;   APPLICANT: Screen, Steven E  
 ;   APPLICANT: Tabaska, Jack E  
 ;   APPLICANT: Cao, Yongwei  
 ;   TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 70131  
 LENGTH: 460  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE: OTHER INFORMATION: Clone ID: MRT4577\_7812C.1.pep  
 US-10-425-114-70131

Query Match Score 45; DB 15; Length 460;  
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 3; Indels 0; Gaps 0;

Qy 3 QALRHQWLHLSFSD 16  
 Db 104 QALDHQHVWVFPD 117

RESULT 10  
 US-10-425-115-559581  
 Sequence 359581, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 359581  
 LENGTH: 83  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE: NAME/KEY: unsure  
 LOCATION: (1)..(83)  
 OTHER INFORMATION: unsure at all Xaa locations  
 US-10-425-115-359581

Query Match Score 44; DB 17; Length 83;  
 Best Local Similarity 58.3%; Pred. No. 31;  
 Matches 7; Conservative 1; Indels 0; Gaps 0;

Qy 6 RIHQWHLHSDF 17  
 Db 53 RSXDWLHVFDDF 64

RESULT 11  
 US-10-425-115-345444  
 Sequence 345444, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 345444 ; LENGTH: 126 ;  
 ; TYPE: PRT ; ORGANISM: Zea mays ;  
 ; FEATURE: OTHER INFORMATION: Clone ID: MRT4577\_7812C.1.pep  
 ; US-10-425-115-345444 ;  
 ; Query Match Score 44; DB 17; Length 126;  
 ; Best Local Similarity 47.1%; Pred. No. 48;  
 ; Matches 8; Conservative 4; Nsmatches 5; Indels 0; Gaps 0;  
 ;  
 ; RESULT 12  
 ; US-10-424-599-232253 ;  
 ; Sequence 232253, Application US/10424599  
 ; Publication No. US2004031072A1 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement.  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 232253  
 ; LENGTH: 173  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE: OTHER INFORMATION: Clone ID: PAT\_MRT3847\_51748C.1.pep  
 ; US-10-424-599-232253 ;  
 ; Query Match Score 44; DB 15; Length 173;  
 ; Best Local Similarity 50.0%; Pred. No. 65;  
 ; Matches 9; Conservative 1; Nsmatches 8; Indels 0; Gaps 0;  
 ;  
 ; RESULT 13  
 ; US-10-424-599-235750 ;  
 ; Sequence 235750, Application US/10424599  
 ; Publication No. US2004031072A1 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 235750  
 ; LENGTH: 184  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE: NAME/KEY: unsure  
 ; LOCATION: (1)..(184)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ;  
 ;

; OTHER INFORMATION: Clone ID: PAT\_MRT3B47\_54910C.1.pep  
 US-10-424-599-233750  
 Query Match Score 44; DB 15; Length 184;  
 Best Local Similarity 63.6%; Pred. No. 69;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 SQAQLRTHWHL 12  
 Db 70 SQALQISENME 80

RESULT 14  
 US-10-408-765A-2745  
 ; Sequence 2745, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Faby, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 ; FILE REFERENCE: 66008.465  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 2745  
 ; LENGTH: 264  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-2745

Query Match Score 44; DB 16; Length 284;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 SQAQLRTHWHL 12  
 Db 155 STGALRIGHNTH 166

RESULT 15  
 US-10-424-599-249784  
 ; Sequence 249784, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovacic, David K  
 ; APPLICANT: Zhou Yinfu  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285584  
 ; SEQ ID NO: 249784  
 ; LENGTH: 475  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE: NAME/KEY: unsure  
 ; LOCATION: (1)...(475)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3B47\_67584C.1.pep  
 US-10-424-599-249784

*This Page Blank (uspto)*

Gencore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## CM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 : Search time 4.45596 Seconds

(without alignments)  
 431.857 Million cell updates/sec

Title: US-10-092-750-40

Perfect score: 106

Sequence: 1 SSQLRITHONLHLSDFDTST 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:

283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79;\*

pir1;\*

pir2;\*

pir3;\*

pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	46	43.4	272	2	S52977		hypothetical prote
2	46	43.4	275	2	D84769		hypothetical prote
3	45	42.5	264	2	T40103		hypothetical prote
4	44	42.5	185	2	A70075		hypothetical prote
5	44	41.5	1276	2	T09204		probable tail-host
6	44	41.5	1291	2	T09273		hypothetical prote
7	43.5	41.0	164	2	D83175		conserved hypothet
8	43	40.6	193	2	A82419		coA-ligase / coenz
9	43	40.6	339	2	E90436		hypothetical prote
10	43	40.6	424	2	C82594		two-component sens
11	43	40.6	443	2	D82975		two-component sens
12	43	40.6	468	2	AD2395		diphthamide synth
13	43	40.6	534	2	S34679		hypothetical prote
14	42	39.6	218	2	H69005		protein-l-isocaspar
15	42	39.6	253	2	A75159		transcription regu
16	42	39.6	292	2	G97325		copper export prot
17	42	39.6	309	2	S52256		hypothetical prote
18	42	39.6	586	2	D82484		SGT protein VCA02
19	42	39.6	750	2	AG3008		polyketide synthet
20	42	39.6	770	2	F98275		safamycin mx1 syn
21	41	38.7	144	2	S69309		hypothetical prote
22	41	38.7	230	2	JC7972		sporomatogenesis-re
23	41	38.7	317	2	H70566		proliferating-cell
24	41	38.7	324	2	F65054		protein C05E11
25	41	38.7	340	2	B88939		hypothetical prote
26	41	38.7	370	2	T09918		hypothetical prote
27	41	38.7	395	2	D64022		conserved hypothet
28	41	38.7	423	2	AH0702		hypothetical prote
29	41	38.7	565	2	S73854		

## ALIGNMENTS

hypothetical prote	30	41	38.7	568	2	T34522
hypothetical prote	31	41	38.7	849	2	T15190
genome polypeptide	32	41	38.7	2127	1	ZLVNSB
genome polypeptide	33	41	38.7	2142	1	ZLVNPV
hypothetical prote	34	41	38.7	5138	2	B96695
hypothetical prote	35	40	37.7	113	2	C95341
probable 5-carboxy	36	40	37.7	116	2	G82048
conserved hypothet	37	40	37.7	181	2	A75341
hypothetical prote	38	40	37.7	210	2	E70772
probable antioxiда	39	40	37.7	212	2	G83215
hypothetical prote	40	40	37.7	214	2	S09896
hypothetical prote	41	40	37.7	235	2	F64799
probable tRNA liga	42	40	37.7	235	2	F90714
probable tRNA liga	43	40	37.7	235	2	C85565
hypothetical prote	44	40	37.7	245	2	T15739
hypothetical prote	45	40	37.7	245	2	AD2269

hypothetical prote	30	41	38.7	568	2	T34522
hypothetical prote	31	41	38.7	849	2	T15190
genome polypeptide	32	41	38.7	2127	1	ZLVNSB
genome polypeptide	33	41	38.7	2142	1	ZLVNPV
hypothetical prote	34	41	38.7	5138	2	B96695
hypothetical prote	35	40	37.7	113	2	C95341
probable 5-carboxy	36	40	37.7	116	2	G82048
conserved hypothet	37	40	37.7	181	2	A75341
hypothetical prote	38	40	37.7	210	2	E70772
probable antioxiда	39	40	37.7	212	2	G83215
hypothetical prote	40	40	37.7	214	2	S09896
hypothetical prote	41	40	37.7	235	2	F64799
probable tRNA liga	42	40	37.7	235	2	F90714
probable tRNA liga	43	40	37.7	235	2	C85565
hypothetical prote	44	40	37.7	245	2	T15739
hypothetical prote	45	40	37.7	245	2	AD2269

hypothetical protein 3 - Erwinia herbicola (fragment)	30	41	38.7	568	2	T34522
Species: Erwinia herbicola	31	41	38.7	849	2	T15190
C:Date: 15-Jul-1995 #sequence_change 01-Sep-1995 #text_change 09-Jul-2004	32	41	38.7	2127	1	ZLVNSB
C:Accession: SS2977	33	41	38.7	2142	1	ZLVNPV
R:Hundie, B.; Alberti, M.; Nivelstein, V.; Beyer, P.; Kleinig, H.; Armstrong, G.A.; Bur: Mol. Gen. Genet. 245, 406-416, 1994	34	41	38.7	5138	2	B96695
A:Title: Functional assignment of Erwinia herbicola Eho 10 carotenoid genes expressed in	35	40	37.7	113	2	C95341
A:Reference number: SS2977	36	40	37.7	116	2	G82048
A:Accession: SS2977	37	40	37.7	181	2	A75341
A:Status: Preliminary	38	40	37.7	210	2	E70772
A:Molecule type: DNA	39	40	37.7	212	2	G83215
A:Cross-references: UNIPROT:Q01334; EMBL:M87280; PIDN:AA649751; PID:914841	40	40	37.7	214	2	S09896

RESULT 1	30	41	38.7	568	2	T34522
SS2977	31	41	38.7	849	2	T15190
hypothetical protein 3 - Erwinia herbicola (fragment)	32	41	38.7	2127	1	ZLVNSB
Species: Erwinia herbicola (house-scar cross)	33	41	38.7	2142	1	ZLVNPV
C:Date: 02-Feb-2001 #sequence_change 02-Feb-2001 #text_change 09-Jul-2004	34	40	37.7	5138	2	B96695
C:Accession: D84769	35	40	37.7	113	2	C95341
R:Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.: Koo, H.; Moffat, K.S.; Cronin, I.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tailor, L.; eau, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.: Nature 402, 761-768, 1999	36	40	37.7	116	2	G82048
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.	37	40	37.7	181	2	A75341
A:Reference number: A84420; PMID:20083487; PMID:10617197	38	40	37.7	210	2	E70772
A:Status: Preliminary	39	40	37.7	212	2	G83215
A:Molecule type: DNA	40	40	37.7	214	2	S09896
A:Cross-references: 1-275 <STO>	41	38.7	568	2	T34522	
C:Genetics:	42	38.7	568	2	T15190	
A:Gene: Atg35500	43	38.7	568	2	ZLVNSB	
A:Map position: 2	44	38.7	568	2	ZLVNPV	
Query Match	45	38.7	568	2	B96695	
Best Local Similarity 43.4%; Pred. No. 6.9%; Matches 8; Conservative 1;	46	38.7	568	2	C95341	
Score 46; DB 2; Length 275;	47	38.7	568	2	G82048	
Query	48	38.7	568	2	A75341	
6 RIHQWHLFSDFT 18	49	38.7	568	2	E70772	

Db 192 RADONRHLYSGFT 204

RESULT 3

T40103 hypothetical protein SPBC2D10\_02 - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: T40103 ; T40105  
 R;Barrell, B.G.; Rajandream, M.A.; Lyne, M.; Skelton, J.; Churcher, C.  
 submitted to the EMBL Data Library, March 1999  
 A;Reference number: Z21005  
 A;Accession: T40103  
 A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA  
 A;Residues: 1-266 <BAR>  
 A;Cross-references: UNIPROT:074796; EMBL:AL049495; PIDN:CAR39854.1; GSPDB:GN00066; SPDB:  
 A;Experimental source: strain 972a-; cosmid C2A9 3p  
 R;Wood, W.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.  
 submitted to the EMBL Data Library, September 1998  
 A;Reference number: 221006  
 A;Accession: T40105  
 A;Status: preliminary; translated from GB/EMBL/DBBJ  
 A;Molecule type: DNA  
 A;Residues: 1-264 <WOO>  
 A;Cross-references: EMBL:AL031788; PIDN:CAK21160.1; GSPDB:GN00067; SPDB:SPBC2D10\_02  
 A;Experimental source: strain 972a-; cosmid c2D10  
 C;Genetics: <BAR1>  
 A;Gene: SPBC2A9\_12  
 A;Map position: 1  
 A;Introns: 61/2  
 C;Genetics: <RQ01>  
 A;Gene: SPBC2D10\_02  
 A;Map position: 2  
 C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC2D10\_02

Query Match 1 SSQALRIHOWLHLFSD 16  
 Best Local Similarity 43.8%; Pred. No. 9.6%; Mismatches 5; Indels 0; Gaps 0;  
 Matches 7; Conservative 5; N mismatches 4; Indels 0; Gaps 0;

Query Match 1 ISSALQIQQMHWIIS 196

RESULT 4

A70075 hypothetical protein YxeG - Bacillus subtilis

C;Species: Bacillus subtilis  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C;Accession: A70075  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-255, 1997  
 A;Authors: Fouger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallei  
 J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hull, M.F.  
 Koeter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lardino, A.; Lardino,  
 A.; Lauber, J.; Lazarevic, V.; Le, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel  
 Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
 A.; Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Saksouka, A.; Seron  
 akechi, M.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tomoda, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yosida, K.  
 A;Authors: Yoonikawa, H.F.; Zumstein, E.; Yoshioka, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A;Reference number: A69780; PMID:98044033; PMID:938377

A;Accession: A70075  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shownA;Molecule type: DNA  
 A;Residues: 1-185 <UNK>

A;Cross-references: UNIPROT:P54946; GB:Z99124; PIDN:92636442; PIDN:GAB15992.

C;Genetics: yxeG

C;Superfamily: Bacillus subtilis hypothetical protein yxeG

Query Match 50 41.5%; Score 44; DB 2; Length 185;  
 Best Local Similarity 41.7%; Pred. No. 9.5%;  
 Matches 10; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

Qy 5 LRIHQWL-----HFSDF 18  
 Db 16 LHIQWIFILWAFLYKHLSFDT 39

RESULT 5

T09204

probable tail-host specificity protein - Streptococcus thermophilus phage Sfi121

C;Species: Streptococcus thermophilus phage Sfi121

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T09204

R;Desiere, F.; Lucchini, S.; Brussow, H.

Virology 241, 345-356, 1998

A;Title: Evolution of Streptococcus thermophilus bacteriophage genomes by modular exchange

A;Reference number: Z16607; PMID:98160788; PMID:949880

A;Accession: T09204

A;Status: translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-1276 &lt;DES&gt;

A;Cross-references: UNIPROT:O64284; EMBL:AF032121; NID:G2935667; PMID:92935667

C;Keywords: tail protein

Query Match 50 41.5%; Score 44; DB 2; Length 1276;  
 Best Local Similarity 53.8%; Pred. No. 76%;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 6 RIHQWLHLSFDT 18

Db 635 RIYQWISLYKEFT 647

RESULT 6

T09273

probable tail-host specificity protein - Streptococcus thermophilus phage Sfi121

C;Species: Streptococcus thermophilus phage Sfi121

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C;Accession: T09273

R;Desiere, F.; Lucchini, S.; Brussow, H.

Virology 241, 345-356, 1998

A;Title: Evolution of Streptococcus thermophilus bacteriophage genomes by modular exchange

A;Reference number: Z16607; PMID:98160788; PMID:9498809

A;Accession: T09273

A;Status: translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-1291 &lt;DES&gt;

A;Cross-references: UNIPROT:O64295; EMBL:AF032122; NID:G2935682; PMID:92935691

Qy 6 RIHQWLHLSFDT 18

Db 635 RIYQWISLYKEFT 647

RESULT 7

D83172 hypothetical protein PA3780 [imported] - Pseudomonas aeruginosa (strain PAO1)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C;Accession: D83172

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen  
A;Reference number: A82950; MUID:2043737; PMID:10984043  
A;Accession: D83172  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-164 <STO>  
A;Cross-references: UNIPROT:Q9HXL5; GB:AE004797; PIDN:9949950; PIDN:AAG0716  
A;Experimental source: strain PAO1  
C;Genetics:  
A;Gene: PA3780

RESULT 8  
A82419 conserved hypothetical protein VCA0769 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1) -  
C;Species: *Vibrio cholerae*  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: A82419  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Pachter, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F., L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A;Reference number: A82035; MUID:2046833; PMID:10952301  
A;Accession: A82419  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-193 <HEI>  
A;Cross-references: UNIPROT:Q9KLH5; GB:AE004405; PIDN:99658186; PIDN:AAF9666  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VCA0769  
A;Map Position: 2  
C;Superfamily: *Escherichia coli* yajB protein

Query Match Score 43.5; DB 2; Length 164;  
Best Local Similarity 41.7%; Pred. No. 10;  
Matches 10; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

Qy	2 SQALRHQWL---HLFSDFTST 20
	:   :   :   :   :   :
Db	84 AQALRLYQWLILALAFFAPAAWST 107

RESULT 9  
E90436 co-enzyme F390 synthetase, probable [imported] - *Sulfolobus solfataricus*  
C;Species: *Sulfolobus solfataricus*  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C;Accession: E90436  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awazy, M.J.; Chan-Pong, T.; Jeffries, A.C.; Koza, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P., arrett, R.A.; Raam, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to Genbank, April 2001  
A;Description: *Sulfolobus solfataricus* complete genome.  
A;Reference number: A99139  
A;Accession: E90436  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-335 <KUR>  
A;Cross-references: UNIPROT:Q97VJ6; GB:AE006641; PIDN:913815936; PIDN:AAK42748.1; GSPDB:CS9022

RESULT 12

AD2395 two-component sensor histidine kinase alr4716 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C;Accession: AD2395

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchishi, N.; Nakazaki, N.; Shimpoo, M.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.; DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AD2395

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-468 <WIE>

A;Cross-references: UNIPROT:Q8YN54; GB:BA000019; PIDN:BAB76415.1; GSPPB:C

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr4716

Query Match 40.6%; Score 43; DB 2; Length 468;

Best Local Similarity 87.5%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 WLHLSDF 17

Db 29 WLHLSDF 36

RESULT 13

S34519 diphthamide synthesis protein DPH2 - yeast (Saccharomyces cerevisiae)

N;A;Alternate names: protein YKL191w

C;Species: Saccharomyces cerevisiae

A;Variety: strain S288C

A;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C;Accession: B59228; S31961; TN0855; S38023; S38024; S52281; S3073; S34679; S49568

R;Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grothues Yeast 9, 1343-1348, 1993

A;Title: Sequencing and analysis of 51.6 kilobases on the left arm of chromosome XI from

A;Reference number: S49688; MUID:94205264; PMID:8154185

A;Accession: B59228

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-534 <WIE>

A;Cross-references: UNIPROT:P32461; EMBL:X74151; PIDN:9450365; PIDN:CAA52247.1; PIDN:g1952

A;Note: the published sequence was revised in GenBank

Yeast 9, 661-667, 1993

R;Cheret, G.; Mattheakis, L.C.; Sor, F.

A;Title: DNA sequence analysis of the YCN2 region of chromosome XI in *Saccharomyces cerevisiae*

A;Reference number: S33960; MUID:93348778; PMID:8394042

A;Accession: S33961

A;Molecule type: DNA

A;Residues: 1-534 <CHE>

A;Cross-references: GB:X69765; PIDN:9296985; PIDN:CAA49420.1; PIDN:g296987

A;Reference number: JN0055; MUID:94010339; PMID:8406038

A;Accession: JN0855

A;Molecule type: DNA

A;Residues: 1-534 <MAT>

RESULT 14

H69005 hypothetical protein MTH1040 - Methanobacterium thermoautotrophicum (strain Delta H)

C;Accession: H69005

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: H69005

R;Qiu, D.; Spadafora, R.; Vitaire, R.; Wang, Y.; Wiezborowski, J.; Gibson, R.; Jiwani, N.; Ji, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Nocella, J.; Reeve, J.N.; J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete Genome sequence of Methanobacterium thermoautotrophicum Delta H: function

A;Reference number: A69000; MUID:98037514; PMID:937163

A;Status: Preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-218 <MTH>

A;Cross-references: UNIPROT:Q27119; GB:AE0000876; GB:AE000666; PIDN:AAB8553

C;Genetic:

A;Gene: MTH1040

A;Start codon: GTG

Query Match 40.6%; Score 43; DB 2; Length 534;

Best Local Similarity 50.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 SSQALRHIQWLHLSDFTST 20

Db 482 SASALONRSWKGSLGSDFDST 501

RESULT 15

A75159 protein-l-isosoaspartate methyltransferase homolog PAB0283 - Pyrococcus abyssi (strain Orse

C;Species: Pyrococcus abyssi

RESULT 16

174 HRWLHVSE 182

C; Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C; Accession: A75159  
R; Anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A; Description: Pyrococcus abyssi Genome sequence: insights into archaeal chromosome number: A75001

A;Accession: A75159  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-253 <KAW>  
A;Cross references: UNIPROT:Q9YJ77; GB:AU248284; GB:AL096836; NID:95457730; PIDN:CAB4935  
A;Experimental source: strain Orsay  
A;Genetic: nmr-like: PDB0283

Query Match 39.6%; Score 42; DB 2; Length 253;  
 Best Local Similarity 41.2%; Pred. No. 28;  
 Marches 7; Conservative 5; Victorian 5; Modern 0.

	matches	$\pi_i$ consecutive	mismatches	$\pi_i$ index	$\pi_i$ caps	$\pi_i$
2y	1	SSQALRTHQWLHLFSDF	17			
2b	102	SGVQVQVQVQVQVQVQVQV	212			

Search completed: November 10, 2004, 13:40:48  
Search time: 5.45506 Secs

This Page Blank (uspto)

Scoring table:	BLOSUM62	Gapop 10.0 , Gapext 0.5	ALIGNMENTS
Searched:	1825181 seqs, 575374646 residues		
Total number of hits satisfying chosen parameters:	1825181		
Minimum DB seq length:	0		
Maximum DB seq length:	20000000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
	Listing first 45 summaries		
Database :	UniProt 02-*		
	1: uniprot_sprot:*		
	2: uniprot_trembl:*		
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
SUMMARIES			
Result No.	Score	Query ID	Description
- - - - -	- - - - -	- - - - -	- - - - -
1	51	48.1	Q9PSZ0 pseudopleur
2	51	48.1	093268 pseudopleur
3	49	46.2	Q9DG39 meleagrid h
4	48	45.3	Q8SVU0
5	48	45.3	420 2 Q8E196 shewaniella
6	47	44.3	Q84SK4 oryza sativ
7	47	44.3	Q7PP27 anopheles g
8	47	44.3	326 2 Q8W029 arabidopsis
9	47	44.3	481 2 Q8PGM7 xanthomonas
10	47	44.3	675 2 Q8GH4
11	47	44.3	675 2 Q9SGH4
12	46	43.4	272 1 YCR3 ERVHE
13	46	43.4	353 2 Q9UD3
14	46	43.4	387 2 Q82290
15	46	43.4	1581 2 Q6BU09
16	46	43.4	2127 2 Q8JTQ9
17	46	43.4	2269 2 Q6T940
18	46	43.4	2269 2 Q6T941
19	46	43.4	2269 2 Q6T942
20	46	43.4	2269 2 AAR87834
21	46	43.4	2269 2 AAR87835
22	46	43.4	2269 2 AAR87836
23	45	42.5	184 2 Q6ZUP8
24	45	42.5	184 2 BAC86171
25	45	42.5	264 2 Q74796
26	45	42.5	474 2 Q8Y308
27	44	41.5	185 1 VXEG_BACSU
28	44	41.5	191 2 Q73QK9
29	44	41.5	191 2 AAS10929
30	44	41.5	229 2 Q7UJM1
31	44	41.5	463 2 Q86HB9
32	44	41.5	503 2 Q9UUT1
33	44	41.5	504 2 Q9C148
34	44	41.5	870 2 Q7RHBS
35	44	41.5	886 2 Q8IJ31
36	44	41.5	910 2 Q37959
37	44	41.5	1070 2 Q8EGF3
38	44	41.5	1276 2 Q64284
39	44	41.5	1291 2 Q64295
40	44	41.5	2671 2 Q754A3
41	44	41.5	2671 2 QAS5340
42	43.5	41.0	164 2 Q9HXL5
43	43.5	41.0	757 1 AOC2_MOUSE
44	43	40.6	92 1 VAPD_NEIGO
45	43	40.6	92 2 Q6ZKQ8
32	44	41.5	Ogunt1 ajellomyces
33	44	41.5	Q9C148 colletotrich
34	44	41.5	Q7RHBS plasmoidium
35	44	41.5	Q8IJ31 lactococcus
36	44	41.5	Q37959 shevannella
37	44	41.5	Q64284 streptococcus
38	44	41.5	Q64295 streptococcus
39	44	41.5	Q754A3 ashbya gossypii
40	44	41.5	QAS5340 ashbya go
41	44	41.5	Q9HXL5 pseudomonas
42	43.5	41.0	Q812C9 mus musculus
43	43	41.0	Q754A3
44	43	40.6	Q6ZKQ8 oryza sativa
45	43	40.6	Q6ZKQ8

RESULT 1  
Q9PSZ0  
ID Q9PSZ0  
AC Q9PSZ0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-Mar-2004 (TREMBLrel. 24, Last annotation update)  
DE Aminopeptidase N (EC 3.4.11.2) (Fragment).  
GN Name=ampN;  
OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes  
americanus);  
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;  
OC Pleuronectoidea; Pleuronectidae; Pseudopleuronectes;  
OX NCBI\_TaxID:8265;  
RN [1] SEQUENCE FROM N.A.  
RP Douglas S.E., Gallant J.W., Bullerwell C.E.;  
RT "Molecular Investigation of Aminopeptidase N Expression in the Winter  
Flounder, Pleuronectes americanus.",  
RA J. Appl. Ichthyol. 0:0-0(1998).  
RL EMBL; AF013383; AAC2807.1; -.  
DR MEROZ; Mol.005; -, -.  
DR GO; GO:0004177; F:aminopeptidase activity; IEA.  
DR GO; GO:0016787; F:hydrolyase activity; IEA.  
DR GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.  
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.  
DR InterPro:IPR01930; Peptidase\_M1.  
DR Pfam:PF01433; Peptidase\_M1\_BS.  
DR PRINS: PR00756; ALADIPASE.  
DR PROSITE; PS00122; ZINC PROTEASE.  
DR Aminopeptidase\_Hydrolase.  
KW FT NON\_TER 680  
SQ SEQUENCE 680 AA; 75883 MN; 0D75D2ECB88EBFC8 CRC64;

;

Query Match 48.1%; Score 51; DB 2; Length 680;  
Best Local Similarity 80.0%; Pred. No. 19;  
Matches 8; Conservative 1; Mismatches 1; Indels 0;  
Gaps 0;

Qy 9 QWLHFSDF18  
Db 182 QWYHLTDFT191

RESULT 2  
O93268  
ID O93268  
AC Q9C148  
DT 01-NOV-1993 (TREMBLrel. 08, Created)  
DT 01-NOV-1993 (TREMBLrel. 08, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Aminopeptidase N (EC 3.4.11.2).

GN	Name=ampN;	DR	InterPro; IPR004340; UI52_0J70.
OS	Pseudopleuronectes americanus (Winter flounder) (Pleuronectes	DR	PF03121; Herpes_UI52 ; 1.
OS	americanus);	DR	Helicase.
OC	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	SQ	SEQUENCE 1078 AA; 120951 MW; AAI1F0CF511BDD65 CRC64;
OC	Acanthomorphia; Neopterygia; Teleostei; Euteleostei; Neorelostei;		
OC	Pleuronectoidii; Acanthopterigii; Percormorpha; Pleuronectiformes;		
OC	Pleuronectoidii; Pleuronectidae; Pseudopleuronectidae;		
NCBI_TaxID=8225;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Intestine;	QY	2 SQALRTHQLHLFSDFT 18
RA	Douglas S.E., Gallant J.W., Bullerwell C.E.;	Db	715 SQSVFARWHTLSEYT 731
RT	"Molecular Investigation of Aminopeptidase N Expression in the Winter		
RT	Flounder. Pleuronectes americanus.";	RESULT 4	
RL	J. Appl. Ichthyol. 0:0 (1998).	Q8SVU0	PRELIMINARY;
DR	MEROPS; NO.001; -;	ID	PRT; 192 AA.
DR	EMBL; AF012465; AAC32794.1; -.	AC	Q8SVU0 ;
DR	GO; GO:0004177; F:aminopeptidase activity; IEA.	DT	01-JUN-2002 (TrEMBLrel. 21, Created)
DR	GO; GO:0016787; F:hydrolase activity; IEA.	DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DR	GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.	DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR	InterPro; IPR001930; Peptidase_M1.	DS	Hypothetical protein ECU04_0880.
DR	InterPro; IPR006028; Pept_M2n_BS.	GN	Name=ECU04_0880;
DR	PFam; PF01433; Peptidase_M1; 1.	OS	Encephalitozoon cuniculi.
DR	PRINTS; PR00716; ALADIPFAPE.	OC	Eukaryota; Fungi; Microsporidia; Encephalitozoon.
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.	NCBI_TaxID=6035;	OX
KW	Aminopeptidase; Hydrolase.	RN	[1]
SQ	SEQUENCE 974 AA;	RN	SEQUENCE FROM N.A.
QY	48 1%; Score 51; DB 2; Length 974;	RC	STRAIN=GB-M1;
Best Local Similarity 80.0%; Pred. No. 27;	Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	RX	MEDLINE=21576510; PubMed=11719806;
DE	UL52 DNA helicase-primease complex protein (UI52 DNA	RA	Katrina M.D., Duprat S., Cornilliat E., Merenier G., Thomarat F.,
DE	helicase-primease complex protein.	RA	Prentier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
DE	Name=HVT00; Synonyms=UI52;	RA	Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
GN	Melagris herpesvirus 1 (herpesvirus of turkeys);	RA	Weissenbach J., Vivares C.P.,
OS	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	RT	"Genome sequence and gene compaction of the eukaryote parasite
OC	Alpha-herpesvirinae; Marek's disease-like viruses.	RL	Encephalitozoon cuniculi."
NCBI_TaxID=37108;	[1]	RN	Nature 414:450-453 (2001).
RN	SEQUENCE FROM N.A.	[2]	RN
RC	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.
Q9DG9Y	PRELIMINARY;	RC	STRAIN=GB-M1;
ID	PRT; 1078 AA.	RX	Genoscope,
AC	Q9DG9Y	RA	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DT	01-MAR-2001 (TrEMBLrel. 16, Created)	RL	EMBL; AL530444; CAD5275.1; -.
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	DR	GO; GO:0004852; F:europorphyrinogen-III synthase activity; IEA.
DT	01-MAR-2004 (TrEMBLrel. 27, Last annotation update)	DR	GO; GO:0006783; P:name biosynthesis; IEA.
DE	UL52 DNA helicase-primease complex protein (UI52 DNA	DR	InterPro; IPR03754; HEN4 synth.
DE	helicase-primease complex protein.	RA	Hypothetical protein.
GN	Name=HVT00; Synonyms=UI52;	SG	SEQUENCE 192 AA; 22306 MW; PAF2E41DD91D0B31 CRC64;
OS	Melagris herpesvirus 1 (herpesvirus of turkeys);	QY	2 SQALRTHQLHLFSD 16
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;	Db	43 SQGIRRERWLHMFT 57
OC	Alpha-herpesvirinae; Marek's disease-like viruses.		
NCBI_TaxID=37108;	[1]	RESULT 5	
RN	SEQUENCE FROM N.A.	O8EI96	PRELIMINARY;
RC	SEQUENCE=FC126;	ID	PRT; 420 AA.
STRAIN=FC126;		AC	Q8EI96;
RX	MEDLINE=21195611; PubMed=11297687;	DT	Q8EI96;
RA	Kingham B.F., Zelnik V., Kopcek J., Majerciak V., Ney E.,	DT	01-MAR-2003 (TrEMBLrel. 23, Created)
RA	"The genome of turkey herpesvirus."	DT	01-OCT-2003 (TrEMBLrel. 23, Last sequence update)
RA	J. Virol. 75:971-978 (2001).	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
RT	Marek's disease viruses."	DE	ATP-dependent RNA helicase_SrmB.
RT	J. Gen. Virol. 82:1123-1135 (2001).	GN	Name=srmB; Order=diocusNames=S00947;
RL	DR	OS	Shewanella oneidensis.
EMBL; AF291866; AAC345790.1; -.	RA	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;	
DR	EMBL; AF222130; AAC30093.1; -.	OC	Alteromonadaceae; Shewanellia.
DR	GO; GO:0003896; F:DNA primase activity; IEA.	NCBI_TaxID=70863;	OX
DR	GO; GO:0004386; F:helicase activity; IEA.	RN	SEQUENCE FROM N.A.
DR	GO; GO:0006260; P:DNA replication; IEA.	RP	



1	RL	Curr. Opin. Plant Biol.	4:447-456 (2001).
[2]	RP	SEQUENCE FROM N.A. MEDLINE=2C249261; PubMed=10785665;	
[3]	RA	SEQUENCE FROM N.A. Stracke R., Werber M., Weisshaar B.; "The WRKY superfamily of plant transcription factors."; Trends Plant Sci. 5:199-206 (2000).	
[4]	RP	SEQUENCE FROM N.A. Jakoby M.J.; Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases. EMBL: AJ419850; CAD12033.1; -.	
[5]	RA	FT NON TER SEQUENCE 326 AA: 34380 MW; 6A60B3C155BB9D9C CRC64;	
QY	1	SSQALRTHQWLH 12 161  : :    1 Db	Query Match 44.3%; Score 47; DB 2; Length 326; Best local Similarity 58.3%; Pred. No. 37; Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Q8PFGM7	Q8PFGM7	PRELIMINARY;	PRT; 481 AA.
AC	0P8GM7;	0P8GM7;	PRELIMINARY;
DT	01-OCT-2002	01-OCT-2002	(TREMBLrel. 22, Created)
DT	01-OCT-2002	01-OCT-2002	(TREMBLrel. 22, Last sequence update)
DE	Integral membrane protein.	Integral membrane protein.	
GN	OrderdLocusNames=EXC3589;	OrderdLocusNames=EXC3589;	
OC	Xanthomonas axonopodis (pv. citri).	Xanthomonas axonopodis (pv. citri).	
OC	Bacteria: Proteobacteria; Gammaproteobacteria; Xanthomonadales;	Bacteria: Proteobacteria; Gammaproteobacteria; Xanthomonadales;	
OC	Xanthomonadaceae; Xanthomonads.	Xanthomonads.	
OX	NCBI_TaxID=928297;	NCBI_TaxID=928297;	
RA	SEQUENCE FROM N.A. STRAIN=306 / ATCC 13902 / XV 101; MEDLINE=2022145C; PubMed=1224217;	SEQUENCE FROM N.A. STRAIN=306 / ATCC 13902 / XV 101; MEDLINE=2022145C; PubMed=1224217;	
RA	da Silva A.C.R.; Ferro J.A.; Reisach F.C.; Farah C.S.; Furulan L.R., Quaggio R.B.; Monteiro-Vitoello C.B.; Van Sluys M.A.; Almeida N.F., Alves L.M.C.; do Amaral A.M.; Bettolini M.C.; Camargo L.E.A., Camarotte G.; Cannavan F.; Cardozo J.; Chambergo F.; Ciapina L.P., Cicarelli J.R.; Coutinho L.L.; Cursino-Santos F.; El-Dorry H., Faria J.B.; Ferreira A.J.S.; Ferreira R.C.C.; Ferro M.I.T., Formighieri E.P.; Franco M.C.; Greggio C.C.; Gruber A., Katsuyama A.M.; Kishi L.T.; Leite R.P.; Lemos E.G.M.; Lemos M.V.F., Locali E.C.; Machado M.A.; Madeira A.M.B.N., Martinuzzi-Rossi N.M., Martins E.C.; Meidanis J.; Menck C.F.M.; Miyaki C.Y.; Moon D.H., Morcira L.M.; Novo M.T.K.; Okura V.K.; Oliveira M.C., Oliveira V.R., Perira H.A.; Rossi A.; Sena J.A.D.; Silva C.; de Souza R.F., Spinola L.A.F.; Takita M.A.; Tamura R.B.; Teixeira E.C.; Tezza R.I.D., Trindade dos Santos M.; Truffi D.; Teei S.M.; White P.F., Setubal J.C.; Kitajima J.P.;		
RA	"Comparison of the genomes of two Xanthomonas pathogens with differing host specificities," Nature 41:459-463 (2002).	"Comparison of the genomes of two Xanthomonas pathogens with differing host specificities," Nature 41:459-463 (2002).	
RA	EMBL: AE012009; AM03843.1; -.	EMBL: AE012009; AM03843.1; -.	
DR	GO; GO:0016020; Camembertae; TEA;	GO; GO:0016020; Camembertae; TEA;	
DR	InterPro; IPR000515; Hypothet_cof.	InterPro; IPR000515; Hypothet_cof.	
DR	InterPro; IPR0040; UBLA; 1;	InterPro; IPR0040; UBLA; 1;	
DR	Pfam; PF01040; UBLA; 1;	Pfam; PF01040; UBLA; 1;	
PROSITE; PS01228; COF_1; UNKNOWN_1.	PROSITE; PS01228; COF_1; UNKNOWN_1.		
KW	Complete proteome.	Complete proteome.	
RN	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,	Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barrister M.E.,	
RA	Submitted (JAN 2001) to the EMBL/GenBank/DDBJ databases.	Submitted (JAN 2001) to the EMBL/GenBank/DDBJ databases.	
RL	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	- - SIMILARITY: Belongs to the bZIP family.	- - SIMILARITY: Belongs to the bZIP family.	
DR	EMBL: BT002502; AA000862.1; -.	EMBL: BT002502; AA000862.1; -.	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	GO; GO:0003677; F:DNA binding; IEA;	GO; GO:0003677; F:DNA binding; IEA;	
DR	DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	DR InterPro; IPR008917; Buk_transcr_DNA.	DR InterPro; IPR008917; Buk_transcr_DNA.	
DR	DR InterPro; IPR048227; TF_B2IP.	DR InterPro; IPR048227; TF_B2IP.	
DR	Pfam; PF00170; bZIP_1.	Pfam; PF00170; bZIP_1.	
DR	SMART; SM00338; BRL2; 1.	SMART; SM00338; BRL2; 1.	
DR	PROSITE; PS50217; B2IP_1.	PROSITE; PS50217; B2IP_1.	
DR	KW DNA-binding; Nuclear Protein.	KW DNA-binding; Nuclear Protein.	
DR	SEQUENCE 675 AA; 953B29B12A762FF85 CRC64;	SEQUENCE 675 AA; 953B29B12A762FF85 CRC64;	
DR	Query Match 44.3%; Score 47; DB 2; Length 675;	Query Match 44.3%; Score 47; DB 2; Length 675;	
DR	Best Local Similarity 58.3%; Pred. No. 82; Gaps 0;	Best Local Similarity 58.3%; Pred. No. 82; Gaps 0;	
DR	Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
DR	YQ 1 SSQALRHQWLH 12  : :    1 Db	YQ 1 SSQALRHQWLH 12  : :    1 510 SSEGKRHQWFH 521.	
DR	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
RA	Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,	Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barrister M.E.,	
RA	Submitted (JAN 2001) to the EMBL/GenBank/DDBJ databases.	Submitted (JAN 2001) to the EMBL/GenBank/DDBJ databases.	
RL	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	- - SIMILARITY: Belongs to the bZIP family.	- - SIMILARITY: Belongs to the bZIP family.	
DR	EMBL: AC011708; AAFL19569.1; -.	EMBL: AC011708; AAFL19569.1; -.	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634; C:nucleus; IEA;	GO; GO:0005634; C:nucleus; IEA;	
DR	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.	
DR	CC 1- SIMILARITY: Belongs to the bZIP family.	CC 1- SIMILARITY: Belongs to the bZIP family.	
DR	NCBI_TaxID=3702;	NCBI_TaxID=3702;	
DR	GO; GO:0005634		

DR GO; GO:0003677; P:DNA binding; IEA.  
 DR GO; GO:0000355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR005917; Euk\_transcr\_DNA.  
 DR InterPro; IPR004827; TF\_BZIP.  
 DR Pfam; PF00170; bZIP; 1.  
 DR SMART; SM00338; BRIZ; 1.  
 DR PROSITE; PS00247; B2IP; 1.  
 DR DNA-Binding; Nuclear Protein.  
 SQ SEQUENCE 675 AA; 73432 MW; DAA74401CS81C078 CRC64;  
 Query Match 44.3%; Score 47; DB 2; Length 675;  
 Best Local Similarity 58.3%; Pred. No. 82;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RN [1] SSQALRTHQWAKH 12  
 Db 510 SSEGKRLHQWPH 521  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N.Bristol N2;  
 RA Mortimore B.;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 YCR3\_ERWHE STANDARD PRT; 272 AA.  
 ID YCR3\_ERWHE  
 AC Q01374;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DT Hypothetical protein in crtE 3' region (CRF3).  
 OS Erwinia herbicola.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Pantoea.  
 NCBI\_TaxID=549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=EBO10;  
 MEDLINE=95107236; PubMed=7808389;  
 RA Hundale B., Albert M., Nevelein V., Beyer P., Kleinig H.,  
 RA Armstrong G.A., Burke D.H., Hearst J.E.,  
 RT "Functional assignment of *Erwinia* herbicola Rhol10 carotenoid genes  
 expressed in *Escherichia coli*";  
 RL Mol. Gen. 245:406-416(1994).  
 CC |- SIMILARITY: Belongs to the sodium:galactoside symporter family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL: MB7280; AAA6975.1; -  
 DR PIR; S52977; S52977.  
 DR InterPro; IPR001927; Na/Gal symport.  
 DR PROSITE; PS000812; NA\_GALACTOSIDE\_SYMP; FALSE\_NEG.  
 KW Hypothetical protein.  
 SQ SEQUENCE 272 AA; 29955 MN; C7C753B416F14AB6 CRC64;  
 Query Match 43.4%; Score 46; DB 1; Length 272;  
 Best Local Similarity 75.0%; Pred. No. 45;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 RN [1] 3 QALRTHQWAKH 14  
 Db 13 QALRTHQWAKH 24  
 RP SEQUENCE FROM N.A.  
 Q9J3D3 PRELIMINARY;  
 ID Q9J3D3  
 AC Q9J3D3;  
 DT 01-MAY-2000 (TRIMBrel. 13, Created)  
 DT 01-MAY-2000 (TRIMBrel. 13, Last annotation update)  
 DT 01-JUN-2003 (TRIMBrel. 24, Last annotation update)  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Mason T.M.,  
 RA Bounseley S.D., Lin X., Kaul S., Shee T.P., Fuji C.Y., Mason T.M.,  
 RA Shen M., Ronning C.M., Fraser C.M., Sonneveld C.R., Venter J.C.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinotsaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Mason T.M.,

```

QY          6 RIHOWHLSFSDP 17
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.K., Miranda M., Narusaka M.,
RA Nguyen M., Onoera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Sato M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.; Submitted (AUG-2001) to the EMBL/GenBank/DDJB databases.
RI          Search completed: November 10, 2004, 13:38:46
RI          Job time : 26.4214 secs
DR          DR GO; GO:0005524; F:ATP binding; IBA.
DR          DR GO; GO:0004765; F:shikimate kinase activity; IEA.
DR          DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR          DR InterPro; IPR00978; HSP0-chsp.
DR          DR InterPro; IPR00623; Shik_Kinase.
DR          DR Pfam; PF01202; SKI; 1.
SQ          SQ SEQUENCE 387 AA; 42687 MW; CF784D442E76503A CRC64;

Query Match        43.4%; Score 46; DB 2; Length 387;
Best Local Similarity 61.5%; Pred. No. 65;           0;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY          6 RIHOWHLSFSDP 18
DR          DR RADQWHLHYSGFT 283

RESULT 15
Q6BUZ0          PRELIMINARY; PRT; 1581 AA.
ID          ID Q6BUZ0
AC          AC
DT          DT 01-OCT-2004 (TREMBLrel. 28, Created)
DT          DT 01-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT          DT 01-Oct-2004 (TREMBLrel. 28, Last annotation update)
Similar to CA1116 | Candida albicans.
ORPNAME=DEHAOC007601G;
GN          OS Debaromyces hansenii (Yeast) (Torulaspora hansenii)
OC          OC Debaromyces; Fungi; Ascomycota; Saccharomycetes;
OC          OC Saccharomycetales; Saccharomycetaceae; Debaromyces.
NCBI_TAXID=4959;
RN          RN [1] _SEQUENCE FROM N.A.
RP          RP STRAIN=CBS767;
RG          RG GENOLAVURES;
RA          RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA          RA Lafontaine I., de Montigny J., March C., Neuveglise C., Talla E.,
RA          RA Goffard N., Frangeul L., Aigle M., Antiochou V., Babour A., Barbe V.,
RA          RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA          RA Boisrame A., Boyer J., Cattolico L., Conaniolieri F., de Darvivar A.,
RA          RA Despous L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA          RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA          RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA          RA Nicard J.M., Nikolski M., Ozta S., Ozier-Kalogeroopoulos O.,
RA          RA Pellenz S., Potier S., Richard G.F., Straub M.J., Suleau A.,
RA          RA Swenney D., Tekka F., Wesołowska-Louvel M., Westhof E., Wirth B.,
RA          RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA          RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA          RA Wincker P., Souciet J.L., "Genomic evolution in yeasts.";
RT          RT Nature 430:35-44 (2004).
RN          RN [2] _SEQUENCE FROM N.A.
RC          RC STRAIN=CBS767;
RA          RA Genoscope;
RA          RA Submitted (JUL-2004) to the EMBL/GenBank/DDJB databases.
RI          RI EMBL; CR382135; CAG86037.1;
SQ          SQ SEQUENCE 1581 AA; 182145 MW; DEPT4AAB263B228 CRC64;

Query Match        43.4%; Score 46; DB 2; Length 1581;
Best Local Similarity 58.3%; Pred. No. 38+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:32:37 ; Search time 6.21762 Seconds  
(without alignments)

191.991 Million cell updates/sec

Title: US-10-092-750-41

Perfect score: 90

Sequence: 1 GQVGRQLAIIGDDINRKR 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cn2\_6\_ptodata/1/iaa/AA-COMB.pep:\*

2: /cn2\_6\_ptodata/1/iaa/5B-COMB.pep:\*

3: /cn2\_6\_ptodata/1/iaa/6A-COMB.pep:\*

4: /cn2\_6\_ptodata/1/iaa/6B-COMB.pep:\*

5: /cn2\_6\_ptodata/1/iaa/FACTUS\_Comb.pep:\*

6: /cn2\_6\_ptodata/1/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	85	94.4	19	3	US-09-236-385A-35		Sequence 35, Appl
2	85	94.4	20	3	US-09-236-385A-36		Sequence 36, Appl
3	85	94.4	27	4	US-09-544-664B-4		Sequence 4, Appl
4	85	94.4	28	1	US-08-440-391-2		Sequence 2, Appl
5	85	94.4	28	1	US-08-440-391-18		Sequence 18, Appl
6	85	94.4	28	2	US-08-908-597A-2		Sequence 2, Appl
7	85	94.4	28	2	US-08-908-597A-18		Sequence 18, Appl
8	85	94.4	28	3	US-09-236-385A-2		Sequence 2, Appl
9	85	94.4	28	3	US-09-236-385A-18		Sequence 18, Appl
10	85	94.4	28	5	PCT-US96-06122-2		Sequence 2, Appl
11	85	94.4	28	5	PCT-US96-06122-18		Sequence 18, Appl
12	85	94.4	36	1	US-08-440-391-14		Sequence 14, Appl
13	85	94.4	36	2	US-08-908-597A-14		Sequence 14, Appl
14	85	94.4	36	3	US-09-236-385A-14		Sequence 14, Appl
15	85	94.4	36	5	PCT-US96-06122-14		Sequence 14, Appl
16	85	94.4	117	4	US-09-381-488-4		Sequence 4, Appl
17	85	94.4	141	1	US-08-471-058-23		Sequence 23, Appl
18	85	94.4	152	1	US-08-471-058-22		Sequence 22, Appl
19	85	94.4	210	3	US-08-471-057-22		Sequence 22, Appl
20	85	94.4	210	4	US-08-470-085-22		Sequence 22, Appl
21	85	94.4	211	1	US-08-321-072A-16		Sequence 7, Appl
22	85	94.4	211	1	US-08-471-058-7		Sequence 7, Appl
23	85	94.4	211	1	US-08-471-058-9		Sequence 9, Appl
24	85	94.4	211	1	US-08-471-058-10		Sequence 10, Appl
25	85	94.4	211	1	US-08-471-058-11		Sequence 11, Appl
26	85	94.4	211	2	US-08-944-530-2		Sequence 2, Appl
27	85	94.4	211	2	US-08-944-530-4		Sequence 4, Appl

RESULT 1  
US-09-236-385A-35  
; Sequence 35, Application US/09236385A  
; Patent No. 6221615  
; GENERAL INFORMATION:  
; APPLICANT: CHITTENDEN, Thomas D.; and  
; LUTZ, Robert J.  
; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
; MODULATE APOPTOSIS  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1415 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: PC DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/236,385A  
; FILING DATE: 25-Jan-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, HENRY N.  
; REGISTRATION NUMBER: 32,073  
; (C) ATTORNEY DOCKET NO. 104322.147CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-942-8400  
; TELEFAX: 202-942-8484  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
; US-09-236-385A-35

Query Match 94.4%; Score 85; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GQVGRQLAIIGDDINR 17  
Db 2 GQVGRQLAIIGDDINR 18

RESULT 2  
US-09-236-385A-36  
Sequence 36, Application US/09236385A  
Patent No. 6221615  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/236,385A  
FILING DATE: 25-Jan-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REGISTRATION NUMBER: 32,073  
(C) ATTORNEY DOCKET NO. 104322.147CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8444  
INFORMATION FOR SEQ ID NO: 36  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 36  
US-09-236-385A-36

Query Match 94.4%; Score 85; DB 3; Length 20;  
Best Local Similarity 100.0%; Prod. No. 1.6e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIIGDDINRR 17  
Db 3 GQVGRQLAIIGDDINRR 19

RESULT 3  
US-09-544-664B-4  
Sequence 4, Application US/09544664B  
GENERAL INFORMATION:  
Patent No. 613280  
APPLICANT: Huang, Ziwei  
APPLICANT: Wang, Jialun  
APPLICANT: Zhang, Zhiqia  
APPLICANT: Shan, Simei  
APPLICANT: Lu, Zhixian  
TITLE OF INVENTION: Enhancement of Peptide Cellular Uptake  
PRIORITY FILING DATE: 2000-04-06  
CURRENT FILING DATE: 2000-04-06  
PRIORITY APPLICATION NUMBER: PCT/US00/09352  
PRIORITY FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: 60/128,202  
PRIOR FILING DATE: 1999-04-07  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 27  
TYPE: PRT

RESULT 4  
US-08-440-391-2  
Sequence 2, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REGISTRATION NUMBER: 32,073  
REFERENCE/DOCKET NUMBER: 104322.147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acid  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-440-391-2

Query Match 94.4%; Score 85; DB 1; Length 28;  
Best Local Similarity 100.0%; Prod. No. 2.3e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQYGRQLAIIGDDINRR 17  
Db 6 GQVGRQLAIIGDDINRR 22

RESULT 5  
US-08-440-391-18  
Sequence 18, Application US/08440391  
Patent No. 5656725  
GENERAL INFORMATION:  
APPLICANT: CHITTENDEN, Thomas D.; and  
LUTZ, Robert J.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

TITLE OF INVENTION: MODULATE APOPTOSIS  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 STREET: Hale and Dorr  
 CITY: Washington Avenue, N.W.  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/440,391  
 FILING DATE: 12-MAY-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 28 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-440-391-18

Query Match 94.4%; Score 85; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIIGDDINRR 17  
 Db 6 GQVGRQLAIIGDDINRR 22

---

RESULT 6  
 US-08-908-597A-2

Sequence 2, Application US/08908597A  
 Patent No. 5863195  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.; and  
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 TITLE OF INVENTION: MODULATE APOPTOSIS  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 STREET: Hale and Dorr  
 CITY: Washington Avenue, N.W.  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/908,597A  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/440,391  
 FILING DATE: 12-MAY-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 REFERENCE/DOCKET NUMBER: 104322.147  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 28 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-908-597A-18

Query Match 94.4%; Score 85; DB 2; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGRQLAIIGDDINRR 17  
 Db 6 GQVGRQLAIIGDDINRR 22

RESULT 8

US-09-236-385A-2  
 / Sequence 2, Application US/09236385A  
 / Patent No. 6221615  
 / GENERAL INFORMATION:  
 / APPLICANT: CHITTENDEN, Thomas D.; and  
 / TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 / NUMBER OF SEQUENCES: 41  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Hale and Dorr  
 / STREET: 1455 Pennsylvania Avenue, N.W.  
 / CITY: Washington  
 / STATE: D.C.  
 / ZIP: 20004  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.25  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/09/236,385A  
 / FILING DATE: 25-Jan-1999  
 / CLASSIFICATION: <Unknown>  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: WIXON, HENRY N.  
 / REGISTRATION NUMBER: 32,073  
 / TELECOMMUNICATION INFORMATION: (C) ATTORNEY DOCKET NO. 104322.147CIP  
 / TELEPHONE: 202-942-8400  
 / TELEFAX: 202-942-8400  
 / INFORMATION FOR SEQ ID NO: 2:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 28 amino acid  
 / TYPE: amino acid  
 / TOPOLOGY: linear  
 / MOLECULE TYPE: peptide  
 / SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 / US-09-236-385A-2  
 / Query Match 94.4%; Score 85; DB 3; Length 28;  
 / Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
 / Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 /  
 / TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 / NUMBER OF SEQUENCES: 41  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Hale and Dorr  
 / STREET: 1455 Pennsylvania Avenue, N.W.  
 / CITY: Washington  
 / STATE: D.C.  
 / ZIP: 20004  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.25  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: PCT/US96/06122  
 / FILING DATE: HERWITH  
 / CLASSIFICATION:  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: US 08/440,391  
 / FILING DATE: 12-MAY-1995  
 / CLASSIFICATION:  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: WIXON, HENRY N.  
 / REGISTRATION NUMBER: 32,073  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: 202-942-8400  
 / TELEFAX: 202-942-8484  
 / INFORMATION FOR SEQ ID NO: 2:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 28 amino acid  
 / TYPE: amino acid  
 / TOPOLOGY: linear  
 / MOLECULE TYPE: peptide  
 / SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 / PCT/US96-06122-2  
 / Query Match 94.4%; Score 85; DB 5; Length 28;  
 / Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
 / Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 GQVGROLAIGDDINRR 22

RESULT 11  
PCT-US96-06122-18  
; Sequence 18, Application PC/TUS9606122

GENERAL INFORMATION:  
APPLICANT: IMMUNOGEN, INC.  
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS  
NUMBER OF INVENTIONS: WHICH MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DCS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06122  
FILING DATE: 12-MAY-1995  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/440,391  
FILING DATE: 12-MAY-1995  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: WIXON, HENRY N.  
REFERENCE/DOCKET NUMBER: 32,073  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

PCT-US96-06122-18

Query Match 94.4%; Score 85; DB 5; Length 28;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;

Qy 1 GQVGROLAIGDDINRR 17

Db 6 GQVGROLAIGDDINRR 22

RESULT 12

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and  
APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
MODULATE APOPTOSIS  
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESS: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 13

US-08-308-597A-14

; Sequence 14, Application US/08908597A

Patent No. 5833795

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 14

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 15

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 16

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 17

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 18

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 19

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 20

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 21

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 22

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 23

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 24

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 25

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 26

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 27

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 28

US-08-440-391-14

; Sequence 14, Application US/08440391

Patent No. 5556725

GENERAL INFORMATION:

APPLICANT: CHITTENDEN, Thomas D.; and

APPLICANT: LUTZ, Robert J.

TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH

SEQUENCE OF INVENTION: MODULATE APOPTOSIS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

RESULT 29

US-08-440-391-14

</div

```

Query Match          94.4%; Score 85; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Number of SEQUENCES: 41

Qy   1 GOVGRQLAIGDDINRR 17
Db    | ||||| | | | | | |
      8 GOVGRQLAIGDDINRR 24

RESULT 14
US-09-216-385A-14
Sequence 14, Application US/09236385A
Patent No. 62215
GENERAL INFORMATION:
APPLICANT: CHITTENDEN, Thomas D.; and
LUTZ, Robert J.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/236,385A
FILING DATE: 25-Jan-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, HENRY N.
REGISTRATION NUMBER: 32,073
(C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-216-385A-14

Query Match          94.4%; Score 85; DB 3; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.1e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Number of SEQUENCES: 34

Qy   1 GOVGRQLAIGDDINRR 17
Db    | ||||| | | | | | |
      8 GOVGRQLAIGDDINRR 24

RESULT 15
PCT-US96-06122-14
Sequence 14, Application PC/TUS9606122
GENERAL INFORMATION:
APPLICANT: IMMUNOGEN, INC.
TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH MODULATE APOPTOSIS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.

```

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 19.3368 Seconds  
(without alignments) 328.807 Million cell updates/sec

Title: US-10-092-750-41

Perfect score: 90

Sequence: 1 GQVGRQLAIGDDINRKR 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgcn2\_6\_ptodata/1/pubpaas/US07\_PUBCOMB\_pep\_\*  
2: /cgna\_6\_ptodata/1/pubpaas/PCT\_NEW\_PUB\_pep\_\*  
3: /cgcn2\_6\_ptodata/1/pubpaas/US05\_NEW\_PUB\_pep\_\*  
4: /cgcn2\_6\_ptodata/1/pubpaas/US06\_PUBCOMB\_pep\_\*  
5: /cgna\_6\_ptodata/1/pubpaas/US07\_NEW\_PUB\_pep\_\*  
6: /cgcn2\_6\_ptodata/1/pubpaas/US08\_NEW\_PUB\_pep\_\*  
7: /cgcn2\_6\_ptodata/1/pubpaas/US09\_PUBCOMB\_pep\_\*  
8: /cgcn2\_6\_ptodata/1/pubpaas/US09A\_PUBCOMB\_pep\_\*  
9: /cgna\_6\_ptodata/1/pubpaas/US09B\_PUBCOMB\_pep\_\*  
10: /cn02\_6\_ptodata/1/pubpaas/US09C\_PUBCOMB\_pep\_\*  
11: /cgcn2\_6\_ptodata/1/pubpaas/US10\_PUBCOMB\_pep\_\*  
12: /cgcn2\_6\_ptodata/1/pubpaas/US10A\_PUBCOMB\_pep\_\*  
13: /cgcn2\_6\_ptodata/1/pubpaas/US10B\_PUBCOMB\_pep\_\*  
14: /cgcn2\_6\_ptodata/1/pubpaas/US10C\_PUBCOMB\_pep\_\*  
15: /cgcn2\_6\_ptodata/1/pubpaas/US10D\_PUBCOMB\_pep\_\*  
16: /cgcn2\_6\_ptodata/1/pubpaas/US11\_PUBCOMB\_pep\_\*  
17: /cgcn2\_6\_ptodata/1/pubpaas/US11\_NEW\_PUB\_pep\_\*  
18: /cgcn2\_6\_ptodata/1/pubpaas/US11\_NEW\_PUB\_pep\_\*  
19: /cgcn2\_6\_ptodata/1/pubpaas/US60\_\_NEW\_PUB\_pep\_\*  
20: /cgcn2\_6\_ptodata/1/pubpaas/US00\_PUBCOMB\_pep\_\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	18	14 US-10-092-750-2	Sequence 2, Appli
2	90	100.0	18	14 US-10-092-750-41	Sequence 35, Appli
3	85	94.4	19	11 US-09-828-870-35	Sequence 36, Appli
4	85	94.4	20	11 US-09-828-870-36	Sequence 2, Appli
5	85	94.4	28	11 US-09-828-870-2	Sequence 18, Appli
6	85	94.4	28	11 US-09-828-870-18	Sequence 14, Appli
7	85	94.4	36	11 US-09-828-870-14	Sequence 0, Appli
8	85	94.4	117	14 US-10-189-294-4	Sequence 22, Appli
9	85	94.4	210	14 US-10-101-482-22	Sequence 7, Appli
10	85	94.4	211	14 US-10-101-482-7	Sequence 9, Appli
11	85	94.4	211	14 US-10-101-482-9	Sequence 10, Appli
12	85	94.4	211	14 US-10-101-482-10	Sequence 11, Appli
13	85	94.4	211	14 US-10-101-482-11	

## ALIGNMENTS

RESULT 1  
US-10-092-750-2  
; Sequence 2, Application US/10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Phillip W.  
; APPLICANT: Alpin, Julia C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; CURRENT APPLICATION NUMBER: US/10/092-750  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US 60/274,525  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: Fast-SEQ for Windows version 4.0  
; SEQ ID NO: 2  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-2

Query Match 100.0%; Score 90; DB 14; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
Matches 18; Conservative 0; Mismatches 0; Indexes 0; Gaps 0;

Qy 1 GQVGRQLAIGDDINRKR 18  
Db 1 GQVGRQLAIGDDINRKR 18

RESULT 2  
US-10-092-750-41  
; Sequence 41, Application US/10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:

APPLICANT: Hammond, Philip W.  
 APPLICANT: Alpin, Julia  
 APPLICANT: Wright, Martin C.  
 TITLE OF INVENTION: Polypeptides Interactive with ECL-X1  
 FILE REFERENCE: 50036/050002  
 CURRENT APPLICATION NUMBER: US10/032,750  
 CURRENT FILING DATE: 2002-03-07  
 PRIOR APPLICATION NUMBER: US 60/274,526  
 NUMBER OF SEQ ID NOS: 253  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 41  
 LENGTH: 18  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-092-750-41

Query Match 100.0%; Score 90; DB 14; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GOVGRQLAIIGDDINRK 18  
 Db 1 GOVGRQLAIIGDDINRK 18

RESULT 3  
 US-09-828-870-35  
 Sequence 35, Application US/09828870  
 Publication No. US20040054129A1  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.; and  
 LUTZ, Robert J.  
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 MODULATE APOPTOSIS  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 APPLICATION NUMBER: US/09/828,870  
 FILING DATE: 10-Apr-2001  
 CLASSIFICATION: <Unknown>  
 CURRENT APPLICATION DATA:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/236,385  
 FILING DATE: 25-JANUARY-1999  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 TELECOMMUNICATION INFORMATION:  
 (C) ATTORNEY DOCKET NO. 104322.147CIP  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 36  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 36  
 US-09-828-870-36

Query Match 94.4%; Score 85; DB 11; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GOVGRQLAIIGDDINRK 17  
 Db 3 GOVGRQLAIIGDDINRK 19

RESULT 5  
 US-09-828-870-2  
 Sequence 2, Application US/09828870  
 Publication No. US20040054129A1  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.; and  
 LUTZ, Robert J.  
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 MODULATE APOPTOSIS  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington

Query Match 94.4%; Score 85; DB 11; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-828-870-35  
 SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
 Query Match 94.4%; Score 85; DB 11; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/828,870  
 FILING DATE: 10-Apr-2001  
 PRIORITY INFORMATION:  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/236,385  
 FILING DATE: 25-JANUARY-1999  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 (C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 28 amino acid  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-828-870-2

RESULT 6  
 US-09-828-870-18  
 Query Match 94.4%; Score 85; DB 11; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GOVGRQLAIGDDINRR 17  
 Db 6 GOVGRQLAIGDDINRR 22

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 14:  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.; and  
 LUTZ, Robert J.  
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/828,870  
 FILING DATE: 10-Apr-2001  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 09/236,385  
 FILING DATE: 25-JANUARY-1999  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 (C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400  
 TELEFAX: 202-942-8484  
 INFORMATION FOR SEQ ID NO: 14:  
 GENERAL INFORMATION:  
 APPLICANT: CHITTENDEN, Thomas D.; and  
 LUTZ, Robert J.  
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: D.C.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/828,870  
 FILING DATE: 10-Apr-2001  
 CLASSIFICATION: <Unknown>  
 APPLICATION NUMBER: 09/236,385  
 FILING DATE: 25-JANUARY-1999  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WIXON, HENRY N.  
 REGISTRATION NUMBER: 32,073  
 (C) ATTORNEY DOCKET NO. 104322.147CIP

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-942-8400

RESULT 8

US-10-189-294-4  
 Sequence 0, Application US/10189294  
 Publication No. US20030060615A1  
 GENERAL INFORMATION:  
 APPLICANT: Kiefer, Michael C.  
     Gibson, Helen L.  
     Fitzpatrick, Paul A.  
 Barr, Phillip J.

TITLE OF INVENTION: A NOVEL Bak BINDING PROTEIN, DNA ENCODING THE PROTEIN, AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWN, THERESA A.  
 STREET: 1560 Broadway, Suite 1200  
 CITY: Denver  
 STATE: CO  
 COUNTRY: USA  
 ZIP: 80202

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/189-294  
 FILING DATE: 01-Jul-2002  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION NUMBER: US/09/381,488  
 FILING DATE: 11-Feb-2000  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BROWN, THERESA A.  
 REGISTRATION NUMBER: 32,547  
 PRIORITY APPLICATION NUMBER: US/09/381,488  
 FILING DATE: 11-Feb-2000  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BROWN, THERESA A.  
 REGISTRATION NUMBER: 4147-15-PUS  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: /note= "Bak (delta) 2 (delta) TM"  
 SEQUENCE CHARACTERISTICS: SEQ ID NO: 4:  
 LENGTH: 117 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..117

US-10-189-294-4  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Query 1 GOVGRQLAIGDDINR 17  
 Db 2 GOVGRQLAIGDDINR 18

RESULT 9  
 US-10-101-482-22  
 Sequence 22, Application US/10101482  
 Publication No. US2003008837A1  
 GENERAL INFORMATION:  
 APPLICANT: KIEFER, MICHAEL C.  
 Barr, Phillip J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto  
 STATE: California

COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/101,482  
 FILING DATE: 18-Mar-2002  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/320,157  
 FILING DATE: 07-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 REFERENCE/DOCKET NUMBER: 23647-20007.20

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 210 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
 US-10-101-182-22

Query Match 94.4%; Score 85; DB 14; Length 210;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GOVGRQLAIGDDINR 17  
 Db 72 GOVGRQLAIGDDINR 88

RESULT 10  
 US-10-101-482-7  
 Sequence 7, Application US/10101482  
 Publication No. US2003008837A1  
 GENERAL INFORMATION:  
 APPLICANT: KIEFER, MICHAEL C.  
 Barr, Phillip J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto  
 STATE: California

COUNTRY: USA  
 ZIP: 94304-1018

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/101,482  
 FILING DATE: 18-Mar-2002  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US/08/320,157  
 FILING DATE: 07-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 REFERENCE/DOCKET NUMBER: 23647-20007.20

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 706141

INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 211 amino acids  
 TYPE: amino acid

TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-101-482-7

Query Match 94.4%; Score 85; DB 14; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
 Matches 17; Conservative 0; N mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGROLAIIIGDDINRR 17  
 Db 72 GQVGROLAIIIGDDINRR 88

RESULT 11  
 US-10-101-482-9

Sequence 9, Application US/10101482  
 Publication No. US20030008837A1

GENERAL INFORMATION:  
 APPLICANT: KIEFER, MICHAEL C.  
 BARR, PHILIP J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
 ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94304-1018

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/101,482  
 FILING DATE: 18-Mar-2002  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION NUMBER: US/08/320,157  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 REFERENCE/DOCKET NUMBER: 23647-20007-20

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 813-5600  
 TELEX: 706141  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 211 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
 US-10-101-482-10

Query Match 94.4%; Score 85; DB 14; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
 Matches 17; Conservative 0; N mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGROLAIIIGDDINRR 17  
 Db 72 GQVGROLAIIIGDDINRR 88

RESULT 13  
 US-10-101-482-11

Sequence 11, Application US/10101482  
 Publication No. US20030008837A1

GENERAL INFORMATION:  
 APPLICANT: KIEFER, MICHAEL C.  
 BARR, PHILIP J.

TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
 ENCODING THE PROTEINS AND METHODS OF USE THEREOF

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORRISON & FOERSTER  
 STREET: 755 Page Mill Road  
 CITY: Palo Alto

US-10-101-482-9

Query Match 94.4%; Score 85; DB 14; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
 Matches 17; Conservative 0; N mismatches 0; Indels 0; Gaps 0;

Qy 1 GQVGROLAIIIGDDINRR 17

STATE: California  
 COUNTRY: USA  
 ZIP: 94304-1018  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/101,482  
 FILING DATE: 18-Mar-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION NUMBER: US/08/320,157  
 FILING DATE: 07-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LEHNHARDT, SUSAN K.  
 REGISTRATION NUMBER: 33,943  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 813-5600  
 TELEFAX: (415) 494-0792  
 TELEX: 706111  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 211 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: Linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
 US-10-101,482-11  
 Query Match 94.4%; Score 85; DB 14; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ 1 GOVGRQLAIGDDINRR 17  
 DB 72 GOVGRQLAIGDDINRR 88

RESULT 14  
 US-10-189-294-2  
 Sequence 2, Application US/1018294  
 Publication No. US2003006015A1  
 GENERAL INFORMATION:  
 APPLICANT: Kieffer, Michael C.  
 Gibson, Helen L.  
 Fitzpatrick, Paul A.  
 Barr, Philip J.  
 TITLE OF INVENTION: A NOVEL BAK BINDING PROTEIN, DNA  
 ENCODING THE PROTEIN, AND METHODS OF USE THEREOF  
 NUMBER OF SEQUENCES: 7  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BROWN, THERESA A.  
 STREET: 1560 Broadway, Suite 1200  
 CITY: Denver  
 STATE: CO  
 ZIP: 80202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/189,294  
 FILING DATE: 01-Jul-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION NUMBER: US/09/381,488  
 FILING DATE: 11-Feb-2000  
 ATTORNEY/AGENT INFORMATION:

NAME: BROWN, THERESA A.  
 REGISTRATION NUMBER: 32,547  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-023  
 INFORMATION FOR SEQ ID NO: 2;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 211 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-10-189-294-2  
 Query Match 94.4%; Score 85; DB 14; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15  
 US-10-177-293-25  
 Sequence 25, Application US/10177293  
 Publication No. US20030124128A1  
 GENERAL INFORMATION:  
 APPLICANT: Glatt, Karen  
 APPLICANT: Zhao, Yumei  
 APPLICANT: Lillie, James  
 APPLICANT: Glatt, Karen  
 APPLICANT: Zhao, Yumei  
 APPLICANT: Gannavarpu, Manjula  
 APPLICANT: Kamatkar, Shubhangi  
 APPLICANT: Mertens, Maureen  
 APPLICANT: Myer, Vic  
 APPLICANT: Wang, Youzhen  
 APPLICANT: Xu, Yongyao  
 APPLICANT: Hoersch, Sebastian  
 APPLICANT: Monahan, John  
 APPLICANT: Meyer, Rachel E.  
 APPLICANT: Bast Jr., Robert C.  
 APPLICANT: Hortobagyi, Gabriel N.  
 APPLICANT: Puszta, Lajos  
 APPLICANT: Meric, Funda  
 APPLICANT: Sahin, Aysegul  
 APPLICANT: Mills, Gordon B.  
 TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
 PREVENTION, AND THERAPY OF BREAST CANCER  
 TITLE OF INVENTION: PREVENTION,  
 PRIORITY: MRE-038  
 CURRENT APPLICATION NUMBER: US/10/177,293  
 CURRENT FILING DATE: 2002-06-21  
 PRIOR APPLICATION NUMBER: US 60/299,887  
 PRIOR FILING DATE: 2001-06-21  
 PRIOR APPLICATION NUMBER: US 60/301,572  
 PRIOR FILING DATE: 2001-06-17  
 PRIOR APPLICATION NUMBER: US 60/306,501  
 PRIOR FILING DATE: 2001-07-18  
 PRIOR APPLICATION NUMBER: US 60/325,002  
 PRIOR FILING DATE: 2001-09-25  
 PRIOR APPLICATION NUMBER: US 60/362,585  
 PRIOR FILING DATE: 2002-03-05  
 PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
 PRIOR FILING DATE: 2002-05-14  
 NUMBER OF SEQ ID NOS: 506  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 25  
 LENGTH: 211  
 TYPE: SRT  
 ORGANISM: Homo sapiens  
 US-10-177-293-25  
 Query Match 94.4%; Score 85; DB 14; Length 211;

Best Local Similarity 100.0%; Pred. No. 5.6e-06;  
Matches 17; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;  
Qy 1 GQVGROLAIGDDINRR 17  
Db 72 GQVGROLAIGDDINRR 88

Search completed: November 11, 2004, 01:28:23  
Job time : 20.3868 secs

This Page Blank (uspto)



RESULT 3  
 AE0564 H+/K+-exchanging ATPase (EC 3.6.3.10) - *Salmonella enterica* subsp. *enterica* serovar Typhi  
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
 A;Note: this species has also been called *Salmonella typhi*  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: AE0564  
 R;Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; Moulé, S.; O'Giora, P.  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi  
 A;Reference number: AB0502; MUID:55875  
 A;Accession: AE0564

Query Match 1 GOVG-ROLAIGDDINR 16  
 Best Local Similarity 51.7%; Pred. No. 6;  
 Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Query Match 1 GOVGWRELAVLGDINK 120  
 Best Local Similarity 51.1%; Pred. No. 19;  
 Matches 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 4  
 AE0564 H+/K+-exchanging ATPase (EC 3.6.3.10) - *Salmonella enterica* subsp. *enterica* serovar Typhi  
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
 A;Note: this species has also been called *Salmonella typhi*  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: AE0564  
 R;Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, S.; Moulé, S.; O'Giora, P.  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi  
 A;Reference number: AB0502; MUID:55875  
 A;Accession: AE0564

Query Match 1 GOVG-ROLAIGDDINR 16  
 Best Local Similarity 51.1%; Pred. No. 19;  
 Matches 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 5  
 CE6479 probable copper-transporting ATPase (EC 3.6.1.-) - *Escherichia coli* (strain K-12)  
 C;Species: *Escherichia coli*  
 C;Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004  
 C;Accession: C6479  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cole, A.; Rose, D.J.; Maurer, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of *Escherichia coli* K-12.  
 A;Reference number: AE64720; MUID:97426617; PMID:9278503  
 A;Accession: C6479  
 A;Molecule type: DNA  
 A;Cross-references: UNIPROT:Q59385; GB:AE000154; GB:U00096; NID:91786683; PIDN: AAC73586.1  
 A;Experimental source: strain K-12, substrate MG1655  
 C;Genetics:

A;Gene: ybar  
 C;Superfamily: *Bacillus* probable copper-transporting ATPase yvgX; ATPase nucleotide-bindin  
 C;Keywords: ATP; copper binding; hydrolase; ion transport; metal binding; phosphoprotein;  
 F:19-38/Domain: heavy-metal-associated homology <TM1>  
 F:105-134/Domain: heavy-metal-associated homology <TM2>  
 F:189-205/Domain: transmembrane #status predicted <TM1>  
 F:218-234/Domain: transmembrane #status predicted <TM2>  
 F:24-58/Domain: ATPase transduction domain homology <ATT>  
 F:438-454/Domain: transmembrane #status predicted <TM3>  
 F:468-484/Domain: transmembrane #status predicted <TM4>  
 F:531-647/Domain: transmembrane #status predicted <TM5>  
 F:643-785/Domain: ATPase nucleotide-binding domain homology <ATN>  
 F:806-822/Domain: transmembrane #status predicted <TM6>  
 F:108-110/113/Binding site: copper (Met, Cys, Cys); #status predicted  
 F:523/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 1 GOVG-ROLAIGDDINR 16  
 Best Local Similarity 51.1%; Pred. No. 19;  
 Matches 3; Mismatches 1; Indels 0; Gaps 0;

Query Match 1 GOVGWRELAVLGDINK 120  
 Best Local Similarity 51.1%; Pred. No. 19;  
 Matches 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 6  
 E88546 probable ATPase ybar [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)  
 C;Species: *Escherichia coli*  
 C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C;Accession: E88546  
 R;Perna, N.T.; Blunkett III, G.; Mau, B.; Glaeser, J.D.; Rose, D.J.; Mayhew, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potatsoussis, K.; Apodaca, Nature 409, 529-533, 2001  
 A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A;Reference number: A85480; MUID:21074935; PMID:11206551  
 A;Accession: E88546  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-833 <PAR>  
 A;Cross-references: UNIPROT:Q8XD24; GB:AE0005174; NID:912513357; PIDN: AAC54833.1; GSPDB: GR  
 C;Genetics:  
 A;Gene: SIV0544  
 C;Superfamily: *Bacillus* probable copper-transporting ATPase yvgX; ATPase nucleotide-bindin

C;Superfamily: Bacillus probable copper-transporting ATPase yyyX; ATPase nucleotide-bind

RESULT 9

Query Match Score 46; DB 2; Length 834;  
Best Local Similarity 51.1%; Pred. No. 19;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GRQLATIGDDIN 15  
Db 712 GRQVANVGDGIN 723

RESULT 7

A90696 Cu(I)-translocation P-type ATPase [imported] - Escherichia coli (strain O157:H7, substrate C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: A90696 R;Yoshizaki, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and Geno  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-334 <HAY>  
A;Cross-references: UNIPROT:QBXD24; GB:BA0000007; PIDN:BAB33960\_1; PID:g13359994; GSPPDB:G  
A;Experimental source: strain O157:H7, substring RIMD 0509952  
C;Genetics:  
A;Gene: Ecs0537

C;Superfamily: Bacillus probable copper-transporting ATPase yyyX; ATPase nucleotide-bind

Query Match Score 46; DB 2; Length 834;  
Best Local Similarity 51.1%; Pred. No. 19;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GRQLATIGDDIN 15  
Db 712 GRQVANVGDGIN 723

RESULT 8

H8:204 cation transport ATPase, E1-E2 family VC2215 [imported] - Vibrio cholerae (strain N16961  
C;Species: Vibrio cholerae  
C;Accession: H8:204 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Dodson, R.J.; Richardson, D.; Brzusto, M.D.; Vaithianathan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellars, E.L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA sequences of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406333; PMID:10953301  
A;Accession: H8:2104  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-915 <HEI>  
A;Cross-references: UNIPROT:Q9KPZ7; GB:AE004293; PIDN:99656766; PIDN:AAF9535  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:

C;Species: Caenorhabditis elegans  
C;Accession: S44824 R;Anderson, K.  
submitted to the EMBL Data Library, September 1993  
A;Description: Sequence of the C. elegans cosmid F54F2.  
A;Accession: S44824  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1226 <AND>  
A;Cross-references: UNIPROT:P34446; EMBL:L23645; PID:9388603; PID:9388605  
C;Genetics:  
A;Introns: 58/2; 137/3; 179/1; 316/2; 393/1; 551/3; 597/2; 662/2; 899/3; 1178/3  
C;Keywords: cytoskeleton; transmembrane protein  
Query Match Score 46; DB 2; Length 1226;  
Best Local Similarity 53.8%; Pred. No. 29;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GOVGRQLATIGDD 13  
Db 359 GVFGQIAVVGDD 371

RESULT 10

S58684 phosphopyruvate hydratase (EC 4.2.1.11) - Helicobacter pylori (strains 26665 and others)

N;Alternate name: enolase  
C;Species: Helicobacter pylori  
C;Date: 29-Nov-1995 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: B64539; S58684 R;Tomb, J.P.; White, O.J.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodet, A.; McKenne, R.; Tomb, J.P.; Kalley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Waithay, L.; Son, J.D.; Nature 38, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C.M.  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394457; PMID:9652185  
A;Accession: B64539  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-426 <TOM>  
A;Cross-references: UNIPROT:PA48285; GB:AE000536; GB:AE000511; PID:g2313230; PIDN:AAD0721:  
A;Experimental source: strain 26695  
R;Schmidt, W.; Odenbreit, S.; Heuermann, D.; Haas, R.  
Mol. Gen. Genet. 248, 563-572, 1995  
A;Title: Cloning of the Helicobacter pylori recA gene and functional characterization of  
A;Reference number: S58683; MUID:96027928; PMID:7476856  
A;Accession: S58684  
A;Molecule type: DNA  
A;Residues: 1-25; I, 27-68 <SCT>  
A;Cross-references: EMBL:235478  
C;Genetic:  
C;Function:  
A;Pathway: glycolysis  
C;Superfamily: enolase  
C;Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase; magnesium  
Pf42/Binding site: magnesium 2 (SER) #status predicted  
P;205, 338/Active site: Glu, Lys #status predicted  
F;242, 286, 313/Binding site: magnesium 1 (ASP, Glu, ASP) #status predicted  
Query Match Score 45; DB 2; Length 426;  
Best Local Similarity 46.2%; Pred. No. 13;  
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 QVGRQLATIGDDIN 15  
Db 786 QQRKTYAMIGDGIN 799

Qy 2 QVGRQLATIGDD 14  
Db 359 GVFGQIAVVGDD 371

Db 303 ELGRQIQLVGDDL 315  
*; Karp, P.; Romero, P.; Zhang, S.*  
*; Science 294, 2317-2313, 2001*  
*A; Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, F.*  
*ster, E.W.*  
*A; Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.*  
*A; Reference number: AB25777; MUID:21608550; PMID:11743193*  
*A; Accession: AC2305*  
*A; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-234 <KTR>*  
*A; Cross-references: UNIPROT:QBUTJF6 ; GB:AE008687 ; PIDN:AA146057\_1 ; PID:g17743817 ; Dupont)*  
*C; Generics:*  
*A; Gene: ori4\**  
*A; Genome: plasmid*

Query Match Score 48.9%; Score 44%; DB 2; Length 234;  
 Best Local Similarity 46.7%; Pred. No. 9.9%;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Qy 3 VGRQLAIIGDDINRR 17  
 Db 12 VGRQKVILGEDVSE 26

RESULT 14  
 JN0527  
*by v-atpase proteolipid PAB1189 - Pyrococcus abyssi (strain Orsay)*  
*C; Species: Pyrococcus abyssi*  
*C; Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004*  
*C; Accession: HT5027*  
*C; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-426 <ARN>*  
*A; Cross-references: UNIPROT:Q9ZMS6 ; GB:AE001453 ; NID:94154651 ; PIDN:AAD0572*  
*A; Experimental source: strain J99*  
*C; Generics:*  
*A; Gene: eno*  
*C; Superfamily: enolase*

Query Match Score 50.0%; Score 45%; DB 2; Length 426;  
 Best Local Similarity 46.2%; Pred. No. 13; Indels 0; Gaps 0;  
 Matches 6; Conservative 6; Mismatches 1; Gaps 0;  
 Qy 2 VGRQLAIIGDDI 14  
 Db 303 ELGRQIQLVGDDL 315

RESULT 15  
 JN0084  
*phytene dehydrogenase (EC 1.3.1.1) - Aphanocapsa sp. (PCC 6714)*  
*N; Alternate names: phytene desaturase*  
*C; Species: Aphanocapsa sp.*  
*C; Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004*  
*C; Accession: JN0084*  
*P; Schmidt, A.; Sandmann, G.*  
*Gene 91, 113-117, 1990*  
*A; Reference number: JN0084; MUID:90382685; PMID:2119316*  
*A; Accession: JN0084*  
*A; Molecule type: DNA*  
*A; Residues: 1-532 <GCH>*  
*A; Cross-references: UNIPROT:P21134 ; GB:M55637 ; NID:g141389 ; PIDN:AAA62573\_1;*  
*A; Note: the authors translated the codon CAG for residue 380 as Gly*  
*C; Comment: This enzyme catalyzes the symmetrical introduction of two double bonds at C-1*  
*C; Genetics:*  
*A; Gene: crtI*  
*C; Superfamily: Aphanocapsa phytoene dehydrogenase*

Query Match Score 49.4%; Score 44.5%; DB 2; Length 532;  
 Best Local Similarity 45.0%; Pred. No. 20; Indels 3; Gaps 1;  
 Matches 9; Conservative 6; Mismatches 2; Gaps 1;

Qy 1 GCGTGTGAI--IGDDINRR 17  
 Db 141 GGTGRLQLIFEGEDVHQ 160

RESULT 13  
 AC3205  
*IS66 family Orf4 (truncated) [imported] - Agrobacterium tumefaciens (strain C58, Dupont)*  
*C; Species: Agrobacterium tumefaciens*  
*C; Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004*  
*C; Accession: AC3205*  
*R; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.*  
*erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, X.; McClell*

Query Match Score 48.9%; Score 44%; DB 2; Length 234;  
 Best Local Similarity 46.7%; Pred. No. 9.9%;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Qy 3 VGRQLAIIGDDINRR 17  
 Db 12 VGRQKVILGEDVSE 26

RESULT 14  
 JN0527  
*by v-atpase proteolipid PAB1189 - Pyrococcus abyssi (strain Orsay)*  
*C; Species: Pyrococcus abyssi*  
*C; Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004*  
*C; Accession: HT5027*  
*C; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-258 <KRW>*  
*A; Cross-references: UNIPROT:Q9UXV1 ; GB:AJ248288 ; GB:AL096836 ; NID:g5458960 ; PIDN:CAB50566;*  
*A; Experimental source: strain Orsay*  
*C; Generics:*  
*A; Gene: PAB1189*  
*A; Reference number: A75001*  
*A; Accession: HT5027*  
*A; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-258 <KRW>*  
*A; Cross-references: UNIPROT:Q9UXV1 ; GB:AJ248288 ; GB:AL096836 ; NID:g5458960 ; PIDN:CAB50566;*  
*A; Experimental source: strain Orsay*  
*C; Generics:*  
*A; Gene: PAB1189*  
*A; Reference number: A75001*  
*A; Accession: HT5027*  
*A; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-258 <KRW>*  
*A; Cross-references: UNIPROT:Q9UXV1 ; GB:AJ248288 ; GB:AL096836 ; NID:g5458960 ; PIDN:CAB50566;*  
*A; Experimental source: strain Orsay*  
*C; Generics:*  
*A; Gene: PAB1189*  
*A; Reference number: A75001*  
*A; Accession: HT5027*  
*A; Status: Preliminary*  
*A; Molecule type: DNA*  
*A; Residues: 1-261 <KTR>*  
*A; Cross-references: UNIPROT:O57733 ; GB:APP000007 ; NID:93236134 ; PIDN:BAA31097\_1 ; PID:g325{*  
*A; Experimental source: strain QT3*  
*A; Note: this accession replaces an interim accession for a sequence replaced by GenBank*  
*C; Generics:*  
*A; Gene: PH970*

Query Match 48.9%; Score 44; DB 2; Length 261;  
Best Local Similarity 43.8%; Pred. No. 11;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 GQVGRGLAIQDDINR 16  
| : | | : | : | : |  
Db 125 GBAGRGFAVVADEIRR 140

Search completed: November 10, 2004, 13:40:49  
Job time : 5.01036 SECs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:27:34 ; Search time 21.9793 Seconds  
(without alignments)

Post-processing: Maximum Match 0%

Listing First 45 summaries

Database : UniProt 02:  
1: uniprot\_sprot:  
2: uniprot\_trembl:

Searched: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181.

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Maximum Match 100%

Listing First 45 summaries

Database : UniProt 02:  
1: uniprot\_sprot:  
2: uniprot\_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	94.4	80	2 O77738	O77738 suis scrofa
2	85	94.4	190	2 Q8NPF3	Q8NPF3 homo sapien
3	85	94.4	211	1 BAK_HUMAN	Q1314 homo sapien
4	85	94.4	211	1 CAG73700	Q1651 homo sapien
5	85	94.4	211	1 Q9M2S6	Cag93700 homo sapien
6	84	93.3	163	2 Q9mz6 evis aries	Q9mz6 ovis aries
7	83	92.2	151	2 Q91WX5	Q91WX5 mus musculus
8	83	92.2	151	2 AAH57589	Aah57589 mus musculus
9	83	92.2	208	1 BAK_MOUSE	Q08264 mus musculus
10	83	92.2	209	2 Q8G264	Q9Jx59 rattus norvegicus
11	83	92.2	209	2 Q9JK59	Q9Jx59 rattus norvegicus
12	49	54.4	336	1 KGP3_BACTIN	Q8az89 bacteroides
13	47	52.2	1430	2 Q8pL12	Q8pL12 xanthomonas
14	46.5	51.7	357	2 Q9Z253	Q9sz53 arabidopsis
15	46	51.1	419	2 Q8AB89	Q8ab89 bacteroides
16	46	51.1	454	2 Q8h716	Q8h716 phytobacter
17	46	51.1	593	2 Q7vbj6	Q7vb16 prochlorococcus
18	46	51.1	832	1 ARCU_SALTY	Q8z854 salmonella
19	46	51.1	832	1 ATCU_ECOLI	Q8zr35 salmonella
20	46	51.1	833	1 ATCU_ECO57	Q8xd24 escherichia
21	46	51.1	833	1 ATCU_ECOLI	Q59385 escherichia
22	46	51.1	834	2 Q7C2W2	Q7C2W2 shigella f1
23	46	51.1	834	2 Q83SE2	Q83SE2 shigella f1
24	46	51.1	834	2 Q8PK77	Q8Fk77 escherichia
25	46	51.1	915	1 ATCU_VIBCH	Q9xp7 vibrio cholerae
26	46	51.1	1226	1 PAT2_CHEEL	P34446 caenorhabditis
27	45.5	50.6	260	2 Q8AXF1	Q89x11 bradyrhizobium
28	45	50.0	211	2 Q89BY1	Q89y11 bradyrhizobium
29	45	50.0	265	2 Q89QG4	Q89qg4 bradyrhizobium
30	45	50.0	297	1 PYRB_HELHP	Q7vit3 helicobacillus
31	45	50.0	424	2 Q82SN5	Q82sn5 nitrosonoma

#### ALIGNMENTS

Q9zms6 helicobacte	32	45	50.0	426	1 ENO_HELHP
E4285 helicobacte	33	45	50.0	426	1 ENO_HELHP
Q82cn7 streptomyce	34	45	583	1	
B21134 synchocystis	35	44.5	49.4	532	1 CRTI_STNY4
Q6cfx2 yarrowia li	36	44.5	49.4	868	2 Q8CPX2
Q8ujv6 agrobacteri	37	44	48.9	234	2 Q8ujv6
Q9uxv1 pyrococcus	38	44	48.9	258	2 Q9UXV1
Q57733 pyrococcus	39	44	48.9	261	2 Q57733
Q7d392 arrobacteri	40	44	48.9	261	2 Q7D392
Q8lef8 arrobacteri	41	44	48.9	329	2 Q8LEF8
Q72b37 desulfobivir	42	44	48.9	395	2 Q72B37
AA96278 desulfococc	43	44	48.9	405	2 AA96278
Q735c2 bacillus ce	44	44	48.9	401	2 Q735c2
Q81bil bacillus ce	45	44	48.9	401	2 Q81BIL
[1] NCBI_TaxID=9823;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darmer D.;					
RL Submitted (JAN-1998) to the Embu/GenBank/DDJB databases.					
DR EMBL: AJ001204; CAA04598; [1]					
DR GO; GO042981; P-regulator of apoptosis; IFA.					
DR InterPro; IPR00712; BCL2_SH.					
DR InterPro; IPR002475; BCL2_Family.					
DR Pfam; PF00452; BCL2; 1.					
DR SMART; SM00337; BCL; 1.					
DR PROSITE; PS50062; BCL2_FAMILY; 1.					
DR PROSITE; PS01259; BH3; 1.					
FT NON_TER [1] 1					
FT NON_TER [80] 80					
SQ SEQUENCE 80 AA; 8818 MN; ED1AFFB3BD7D59C86 CRC64;					
Query Match 94.4%; Score 85; DB 2; Length 80;					
Best Local Similarity 100.0%; Pred. No. 9.9e-06;					
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 1 GOYGRQLATIGDDINRR 17					
Db 23 GQYGRQLATIGDDINRR 39					
RESULT 2					
Q8NPF3 PRELIMINARY; PRT; 190 AA.					
ID Q8NPF3; Q8NPF3; [1]					
AC Q8NPF3; [1]					
DT 01-OCT-2002 (TrEMBLrel. 22, Created)					
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)					
DE Pro-apoptotic protein BAK variant.					
GN Name=BAK;					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI_TaxID=96006; [1]					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Ma J.;					

RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.  
 EMBL; AF520590; AACM74949.1.; -.  
 HSSP; Q16611; 1BXL.  
 DR GO; GO:004281; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2 BH.  
 DR InterPro; IPR002475; BCL2\_-family.  
 DR PFAM; PF00452; BC1-2; 1.  
 SMART; SM00337; BCL; 1.  
 PROSITE; PS01080; BH1; 1.  
 PROSITE; PS01258; BH2; 1.  
 PROSITE; PS01259; BH3; 1.  
 SQ SEQUENCE 190 AA; 21231 MW; A9D4EB8526D0897B CRC64;  
 Query Match 94.4%; Score 85; DB 2; Length 190;  
 Best Local Similarity 100.0%; Pred No 2.4e-05;  
 Matches 17; Conservative 0; Mismatches 0; Gaps 0;  
 Indels 0;  
 Qy 1 GOVGRQLAIGDDINRR 17  
 Db 51 GOVGRQLAIGDDINRR 67

RESULT 3  
 BAK2\_HUMAN STANDARD; PRT; 211 AA.  
 ID BAK2\_HUMAN  
 AC Q13014;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE Putative Bcl-2 homologous antagonist/killer 2 (apoptosis regulator  
 DE BAK2).  
 OS Homo sapiens (Human).  
 OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92231664; PubMed=7715731;  
 RA Tomei L.D., Brauer M.J., Powers V.C., Wu J.J., Umansky S.R.,  
 RT Bak;"  
 RL Nature 374:736-739 (1995).  
 CC --!- FUNCTION: In the presence of an appropriate stimulus, accelerates  
 CC programmed cell death by binding to, and antagonizing the a  
 CC repressor Bcl-2 or its adenovirus homolog E1B 19k protein.  
 CC --!- SUBCELLULAR LOCATION: Membrane-associated (Potential).  
 CC --!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues, with  
 CC highest levels in the heart and skeletal muscle.  
 CC --!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and  
 CC BAX for their pro-apoptotic activity and for their interaction  
 CC with anti-apoptotic members of the Bcl-2 family. Apoptotic members  
 CC of the Bcl-2 family.  
 CC --!- SIMILARITY: Belongs to the Bcl-2 family.  
 CC --!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
 CC --!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
 CC --!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.  
 CC --!- CAUTION: This is probably the product of a pseudogene.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 DR EMBL; U16812; AAA71467.1; -.  
 DR PIR; S58875; S58875; -.

DR HSSP; Q16611; 1BXL.  
 DR Genew; HGNC:996; BCL2L7P1.  
 DR GO; GO:0016020; C:membrane; NAS.  
 DR InterPro; IPR00712; Bcl2 BH.  
 DR InterPro; IPR002475; BCL2\_-family.  
 DR Pfam; PF00452; BC1-2; 1.  
 SMART; SM00337; BCL; 1.  
 PROSITE; PS00062; BCL2\_-FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 KW Apoptosis; Hypothetical protein; Transmembrane.  
 FT DOMAIN 74 88 BH3.  
 FT DOMAIN 117 136 BH1.  
 FT DOMAIN 169 184 BH2.  
 FT TRANSEM 188 205 Potential.  
 SQ SEQUENCE 211 AA; 23411 MW; 703875EC4DCCC1D3 CRC64;  
 Query Match 94.4%; Score 85; DB 1; Length 211;  
 Best Local Similarity 100.0%; Pred No 2.6e-05;  
 Matches 17; Conservative 0; Mismatches 0; Gaps 0;  
 Indels 0;  
 Gaps 0;

RESULT 4  
 BAK\_HUMAN STANDARD; PRT; 211 AA.  
 ID BAK\_HUMAN  
 AC Q16611; Q92533; STANDARD; PRT; 211 AA.  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE Bcl-2 homologous antagonist/killer (Apoptosis  
 DE like 7 protein).  
 DE Name=BAK1; Synonyms=BAK, BCL2L7;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 NCBI\_TaxID=9606/  
 RN Sequence FROM N.A.  
 RC TISSUE-B-cell;  
 RX MEDLINE=95231652; PubMed=7715729;  
 RA Farrow S.N., White J.H.M., Martinou J., Raven T., Pun K.-T.,  
 RA Grinham C.J., Martinou J.C., Brown R.;  
 RA "Cloning of a bcl-2 homologue by interaction with adenovirus E1B  
 RT 19K.";  
 RT Nature 374:733-739 (1995).  
 RN [2]  
 RP Sequence FROM N.A.  
 RX MEDLINE=9531655; PubMed=7715730;  
 RA Chittenden T., Harrington E.A., O'Connor R., Flemington C., Lutz R.J.,  
 RA Evan G.I., Guild B.C.;  
 RA "Induction of apoptosis by the Bcl-2 homologue Bak.";  
 RT Nature 374:733-736 (1995).  
 RN [3]  
 RP Sequence FROM N.A.  
 RX MEDLINE=9531654; PubMed=7715731;  
 RA Kiefer M.C., Brauer M.J., Powers V.C., Wu J.J., Umansky S.R.,  
 RA Tomei L.D., Barr P.J.;  
 RA "Modulation of apoptosis by the widely distributed Bcl-2 homologue  
 RT Bak.";  
 RT Nature 374:736-739 (1995).  
 RN [4]  
 RP Sequence FROM N.A. AND VARIANTS VAL-28 AND ARG-69.  
 RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,  
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
 RA Schachwitz W.S., Sherwood J.K., Witruk L.A., Nickerson D.A.;  
 RA "NIH-SVs, environmental genome project, NIEHS ES15478, Department  
 of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu/>)."

Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

[5] SEQUENCE FROM N.A.; PubMed=14574404; DOI=10.1038/nature02055;

RP Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L., Wilming J., Jones M.C., Horton R., Hunt S.E., Scott C.E., Gilbert J.G.R., Clamp M.B., Bethel G., Milne S., Ainscough R., Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S., Babbage A.K., Baggley C.L., Bailey J., Banerjee R., Barker D.J., Barlow K.F., Bates K., Beare D.M., Besley H., Beasley O., Bird C.P., Blayke S.E., Bray-Alien S., Brook J., Brown J.Y., Burford D.C., Burhill W., Burton J., Gardner C., Carter N.P., Chapman J.C., Clark S.Y., Clark G., Cleee C.M., Clegg S., Cobley V., Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J., Culley K.M., Dhami P., Davies J., Dunn M., Earle-Brown M.E., Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A., Franklin J., French L., Garner P., Ghori M.J., Gilby L.M., Gitterl J., Githero R.J., Garnett J., Ghori M.J., Gribble S., Griffiths C., Griffths M.N.D., Hall R., Hallis K.S., Hammond S., Harley J.L., Hart P.D., Heathcott R., Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Hucke B., Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A., Lawlor S., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C., Loveall J.E., Lovell J., Martin S., Mashraghi-Mohammadi M., McMurray A., Moore M.J.F., Mullikin J.C., McLaren S.J., McLay K., Novik K.L., Oliver K., Overton-Larby E.K.C., Parker A., Patel R., Pearce A.V., Peck A.I., Phillipine B.J.C.T., Phillips S., Plumb R.W., Porter K.M., Ramsey Y., Randy S.A., Rice C.M., Ross M.T., Seale S.M., Sehra H.K., Sherrard E., Skuse C.D., Smith S.M., Spraggon L., Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J., Thaker A.J., Thomas D.W., Thorpe A., Tracey A., Tronians A., Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.D., Whittaker H., Wild A., Willsey D.J., Wilmer T.E., Wood J.M., Wray P.W., Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A., Durbin R., Hubbard T., Sulston J.J., Dunham I., Rogers J., Beck S., RT "The DNA sequence and analysis of human chromosome 6.";

RN Nature 425:805-811(2003).  
 RN Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 RN [5] SEQUENCE FROM N.A.; PubMed=14574404; DOI=10.1038/nature02055;  
 RP Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L., Wilming J., Jones M.C., Horton R., Hunt S.E., Scott C.E., Gilbert J.G.R., Clamp M.B., Bethel G., Milne S., Ainscough R., Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S., Babbage A.K., Baggley C.L., Bailey J., Banerjee R., Barker D.J., Barlow K.F., Bates K., Beare D.M., Besley H., Beasley O., Bird C.P., Blayke S.E., Bray-Alien S., Brook J., Brown J.Y., Burford D.C., Burhill W., Burton J., Gardner C., Carter N.P., Chapman J.C., Clark S.Y., Clark G., Cleee C.M., Clegg S., Cobley V., Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J., Culley K.M., Dhami P., Davies J., Dunn M., Earle-Brown M.E., Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A., Franklin J., French L., Garner P., Ghori M.J., Ghori M.J., Gilby L.M., Gitterl J., Githero R.J., Garnett J., Ghori M.J., Gribble S., Griffiths C., Griffths M.N.D., Hall R., Hallis K.S., Hammond S., Harley J.L., Hart P.D., Heathcott R., Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Hucke B., Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A., Lawlor S., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C., Loveall J.E., Lovell J., Martin S., Mashraghi-Mohammadi M., McMurray A., Moore M.J.F., Mullikin J.C., McLaren S.J., McLay K., Novik K.L., Oliver K., Overton-Larby E.K.C., Parker A., Patel R., Pearce A.V., Peck A.I., Phillipine B.J.C.T., Phillips S., Plumb R.W., Porter K.M., Ramsey Y., Randy S.A., Rice C.M., Ross M.T., Seale S.M., Sehra H.K., Sherrard E., Skuse C.D., Smith S.M., Spraggon L., Squares S.L., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J., Thaker A.J., Thomas D.W., Thorpe A., Tracey A., Tronians A., Wall M., Wallis J.M., West A.P., White S.S., Whitehead S.D., Whittaker H., Wild A., Willsey D.J., Wilmer T.E., Wood J.M., Wray P.W., Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A., Durbin R., Hubbard T., Sulston J.J., Dunham I., Rogers J., Beck S., RT "The DNA sequence and analysis of human chromosome 6.";

RN Nature 425:805-811(2003).  
 RN [6] SEQUENCE FROM N.A.; PubMed=14574404; DOI=10.1038/nature02055;

RC TISSUE=Lung; PubMed=14574404; DOI=10.1038/nature02055;

RX Strasberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Tester J., Klausner R.D., Collins F.S., Wagner L., Sheppard C.M., Schulter G.D., Altshuler S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhattacharyya P.H., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marsilia K., Farmer P., Casavant T.L., Scheetz T.E., Stapanian M., Soares M.B., Ronald M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S., Loqueland N.A., Peters G.J., Abramson R.D., Mallahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pahay J., Helton E., Ketterman M., Madan A., Rodriques S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Schmitz J., Myers R.M., Rodriguez A.C., Grimwood J., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smialius D.E., Schnetzer A., Schein J.E., Jones S.J.M., Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [7] SEQUENCE OF 96-206 FROM N.A.

RP Eguchi H., Hayashi S.; RT "Estrogen alters expression of apoptosis-regulators, Bcl-2, Bcl-xL and cancer cells.";

RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RN [8] MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.

RX MEDLINE=96091131; PubMed=8521816;

RN Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J., Ebb R.J.; RT "A conserved domain in Bak, distinct from BH1 and BH2, mediates cell death and protein binding functions.";

RX Eblangovan B., Chinmadrula G., Lutz R.J.; RT

RX "BH3 domain is required for BH1 and BH2 mediated cell death and protein binding functions.";

RX EMBL J. 14:5589-5596(1998).  
 RN [9] STRUCTURE BY NMR OF 72-87.

RX Sattler M., Liang H., Nettesheim D., Meadows R.P., Harlan J.E., Eberstadt M., Yoon H.S., Shuker S.B., Chang B.S., Minn A.J., Thompson C.B., Pesik S.W.; RT "Structure of Bcl-xL-Bak peptide complex: recognition between RT regulators of apoptosis";

RL CC FUNCTION: In the presence of an appropriate stimulus, accelerates programmed cell death by binding to, and antagonizing the a repressor Bcl-2 or its adenovirus homolog E1B 19k protein.

CC CC SUBUNIT: Forms heterodimers with Bcl-2, E1B 19k protein, and Bcl-X(L).

CC CC SUBCELLULAR LOCATION: Membrane-bound (Potential).

CC CC TISSUE SPECIFICITY: Expressed in a wide variety of tissues, with highest levels in the heart and skeletal muscle.

CC CC DOMAIN: Intract BAK, BAD and BAX for their pro-apoptotic activity and for their interaction with anti-apoptotic members of the Bcl-2 family.

CC CC SIMILARITY: Belongs to the Bcl-2 family.

CC CC SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.

CC CC SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.

CC CC SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC DR EMBL: X84213; CAA58997.1; -.

DR EMBL: U3765; AAA93066.1; -.

DR EMBL: U16811; AAJ74466.1; -.

DR EMBL: AY260471; AAO74828.1; -.

DR EMBL: 293017; CAB65626.1; -.

DR EMBL: BC004411; AAH04431.1; -.

DR EMBL: D88397; BAA13606.1; -.

DR EMBL: D88396; BAA13605.1; JOINED.

DR PIR: SS8873; SS8873.

DR PIR: 18XL; MNR: B=72-87.

DR GeneID: HGNC:949; BAK1.

DR MIN: 600516; -.

DR GO:GO008337; P:apoptotic mitochondrial changes; TAS.

DR InterPro: IPR000712; BCL2\_BH.

DR InterPro: IPR02475; BCL2\_FAMILY.

DR PFAM: PF00455; Bcl-2\_1.

DR PROSITE: PS00062; BCL2\_FAMILY; 1.

DR PROSITE: PS01080; BH1\_1.

DR PROSITE: PS01258; BH2\_1.

DR PROSITE: PS01259; BH3\_1.

KW 3D-structure: Apoptosis; Polymorphism; Transmembrane.

FT DOMAIN: 74 88 BH3.

FT DOMAIN: 117 136 BH1.

FT DOMAIN: 169 184 BH2.

FT TRANSMEM 188 205 Potential.

FT VARIANT 28 28 A->V.

FT VARIANT /FTId=VAR\_018829.

FT VARIANT 69 69 S->R.

FT VARIANT /FTId=VAR\_018830.

FT TURN 74 75

FT HELIX 76 84

FT TURN 85 85

FT SEQUENCE 211 AA; 23409 MW; A2200DE72AA6DD4E CRC64;

SQ Query Match 94.4%; Score 85; DB 1; Length 211;



Qy	1 GOVGRQLAIIQDDINRR 17	RC STRAIN=Swiss; TISSUE=Liver;
Db	70 GOVGRQLAIIQDDINRR 86	RX MEDLINE:974-6138; PubMed=299236;
		RA Ulrich E.; Kauffmann-Zeh A.; Hueber A.O.; Williamson J.,
		RA Chittenden T.; Ma A.; Evan G.I.;
		RT "Gene structure, cDNA sequence, and expression of murine Bak, a proapoptotic Bcl-2 family member.";
		RT RLI Genomics 44:195-200(1997).
		-1- FUNCTION: In the presence of an appropriate stimulus, accelerates programmed cell death by binding to, and antagonizing the a repressor Bcl-2 or its adenovirus homolog E1B 19k protein (By similarity).
		-1- SUBUNIT: Forms heterodimers with Bcl-2, B1B 19k protein, and Bcl-2B.
		-1- X(L) (By similarity)
		-1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
		-1- TISSUE SPECIFICITY: Widely expressed.
		-1- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and BAX for their pro-apoptotic activity and for their interaction with anti-apoptotic members of the Bcl-2 family (By similarity).
		-1- SIMILARITY: Belongs to the Bcl-2 family.
		-1- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.
		-1- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
		-1- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
		CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage of by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licensee@isb-sib.ch).
		CC EMBL: Y13231; CDA:73684.1; -.
		CC RX MEDLINE:2238257; PubMed:1247932;
		RA Klaunzer R.D., Collins F.S., Wagner L.H., Derge J.G., Feingold E.A., Grouse L.H., Scheuer G.D., Schaefer C.P., Bhat N.K., Zeeberg B., Bustow K.H., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F., Stapleton M., Soares M.B., Bonaldo M.F., Cesavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullally S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S.W., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kertemian M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Bouffard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Rodriguez A.C., Grimwood J., Skalska U., Smailius D.E., Scherfield Y.S., Krzywinski M.I., Jones S.J., Matra M.A., Scherfield A., Schein J.U., RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; RT and mouse cDNA sequences"; RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [2]
		RP SEQUENCE FROM N.A.
		RC STRAIN=CSPL/6; TISSUE=Brain;
		RA Strasbourg R.; Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
		RL DR EMBL: BC057589; AAH57589.1; -.
		SEQUENCE 151 AA; 16402 MW; 18C13BFFF86E4F33B CRC64;
		Query Match 92.2%; Score 83; DB 2; Length 151; Best Local Similarity 94.1%; Pred. No. 3.9e-05; Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0; Gaps 0;
Qy	1 GOVGRQLAIIQDDINRR 17	Query Match 92.2%; Score 83; DB 1; Length 208; Best Local Similarity 94.1%; Pred. No. 5.4e-05; Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0; Gaps 0;
Db	70 GOVGRQLAIIQDDINRR 86	Qy 1 GOVGRQLAIIQDDINRR 17
		Db 69 GQvGRQLAIIQDDINRR 85
		RESULT 10
		BAK_MOUSE ID Q8C264 STANDARD; PRT; 208 AA.
		AC 008734; DT 01-NOV-1997 (Rel. 35, Created)
		DT 01-NOV-1997 (Rel. 35, Last sequence update)
		DT 05-JUL-2004 (Rel. 44, Last annotation update)
		DE Bcl-2 homologous antagonist/killer (Apoptosis regulator BAK).
		GN Name=Bak; Synonyms=Bak;
		OS Mus musculus (Mouse)
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
		NCBI_TaxID=10090; RN [1]
		RP SEQUENCE FROM N.A.



Science 299:2074-2076(2003).  
 CC -!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-  
 CC fructose 1,6-bisphosphate.  
 CC -!- PATHWAY: Key control step of glycolysis.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the phosphofructokinase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC purposes requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; AE016940; AAO78462.1;  
 DR HSSP; P06938; 2PFK.  
 DR HAMAP; MF\_00339; -;  
 DR InterPro; IPR000023; Ppfckinase.  
 DR PFam; PF00365; Pfk; 1.  
 DR PRINTS; PR00476; PHFRCTKINASE.  
 DR ProDom; P0000707; Ppfckinase; 1.  
 DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; FALSE NEG.  
 DR KW Allosteric enzyme; ATP-binding; Complete proteome; Glycolysis; Kinase;  
 KW Magnesium; Transferase.  
 FT NP\_BIND 20 24 ATP (By similarity).  
 FT NP\_BIND 160 164 ATP (By similarity).  
 FT NP\_BIND 177 193 ATP (By similarity).  
 FT ACT\_SITE 133 133 Proton acceptor (By similarity).  
 FT BINDING 168 168 Substrate (By similarity).  
 FT BINDING 255 255 Substrate (By similarity).  
 FT BINDING 264 264 Substrate (By similarity).  
 SQ SEQUENCE 336 AA; 990BAE6B5BD5F79 CRC64;  
 SQ  
 Query Match 54.4%; Score 49; DB 1; Length 336;  
 Best Local Similarity 60.0%; Pred. No. 24;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 GQVGRQLAIIQDDINR 15  
 Db 290 GQFGRMIALRGDDIS 304  
 RESULT 13  
 Q8PLI2 PRELIMINARY; PRT; 1430 AA.  
 AC 08PLI2; PRELIMINARY;  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hemagglutinin/hemolysin-related protein.  
 GN OrderедLocusName=ZAC1816;  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacterium; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OC NCBI\_TaxID=92629;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 MEDLINE=2202245; PubMed=12024217;  
 RX da Silva A.C.R., Farro J.A., Reinach F.C., Farach C.S., Furian L.R.,  
 RA Quaggio R.B., Monteiro-Vitoello C.B., Van Sluys M.A., Almeida N.P.,  
 RA Alves L.M.C., do Amaral A.M., Berrolini M.C., Camargo L.B.A.,  
 RA Camarotte G., Cannavan F., Cardozo J., Chambengo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Gregorio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado W.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA MeiaC.C., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RN  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.,  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 host specificities."  
 RL Nature 417:459-463 (2002);  
 DR EMBL; AE01815; AAM36578.1;  
 DR InterPro; IPR00819; Fil\_haemagg.  
 DR Pfam; PF05584; Fil\_haemagg; 5.  
 KW Complete proteome.  
 SEQUENCE 1430 AA; 145025 MW; 6BA2BCD1BEB27E053 CRC64;  
 Query Match 52.2%; Score 47;  
 Best Local Similarity 52.9%; Pred. No. 2.2e+02;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 GQVGRQLAIIQDDINR 17  
 Db 1218 GQVGRQLAIIQDDINRQ 1234  
 RESULT 14  
 Q9SZ53 PRELIMINARY; FRT; 357 AA.  
 AC Q9SZ53;  
 ID Q9SZ53; PRELIMINARY;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Protein phosphatase 2C-like protein (ATR3186/F11C18\_60).  
 GN Name=ATR4931860;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Villarcelo R., Gielens J., van Montagu M., Hohenstein J., Newes H.W.,  
 RA Lemke K., Mayer K.F.X.;  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Becker J.R.;  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G.,  
 RA Shinozaki K., Davis R.W., Theologis A., Becker J.R.;  
 RA Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G.,  
 RA Shinozaki K., Davis R.W., Theologis A., Becker J.R.;  
 RA Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G.,  
 RA Shinozaki K., Davis R.W., Theologis A., Becker J.R.;  
 RA Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Terryn N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,  
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villarcelo R.,  
 RA Gielens J., Van Montagu M., Newes H.W., Lemcke K., Mayer K.F.X.;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 DR AL049607; CAB40756.1; -.

DR EMBL; AY057611; AAU14406\_1; -;  
 DR EMBL; AY113024; AAU47332\_1; -;  
 DR EMBL; ALI61579; CAE79904\_1; -;  
 DR PIR; T06308; T06308.  
 DR HSSP; P35113; 1A6Q.  
 DR GO; GO:003824; F: catalytic activity; EA.  
 DR InterPro; IPR000222; PPC.  
 DR InterPro; IPR001132; PPC-like.  
 DR Pfam; PF00481; PP2C.  
 DR SMART; SM00332; PP2C.  
 DR SMART; SM00331; PP2C SIG.  
 DR PROSITE; PS0102; PP2C; UNKNOWN\_1.  
 SQ SEQUENCE 357 AA; 39203 MW; 98EE1A0981BCC0D3 CRC64;  
 Query Match 51.7%; Score 46.5%; DB 2; Length 357;  
 Best Local Similarity 58.8%; Pred. No. 65;  
 Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
 Qy 1 GQVG-RQLAIGDDNIR 16  
 Db 104 GQRGMRELAVLGDKINK 120

---

RESULT 15

Q8ABB9	PRELIMINARY;	PRT;	419 AA.
ID Q8ABB9;			
AC Q8ABB9;			
DT 01-JUN-2003	(T)EMBLrel.	24;	Created
DT 01-JUN-2003	(T)EMBLrel.	24;	Last sequence update
DT 01-JUN-2003	(T)EMBLrel.	24;	Last annotation update
DE Hypothetical protein.			
DR OrderedLocusNames=BP0221;			
GN Bacteroides thetaiotaomicron.			
OS Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;			
OC Bacteroidaceae; Bacteroides.			
OC NCBI TaxID=818;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=VPI-5482 / ATCC 29148;			
RX MEDLINE=22550888; PubMed=1263928;			
RA Xu J., Bjursell M.K., Hirnrod J.J., Deng S., Carmichael L.K.,			
RA Chiang H.C., Hooper L.V., Gordon J.I.; Gordon J.I.;			
RT "A genomic view of the human-Bacteroides thetaiotomicron symbiosis.";			
RL Science 299:2074-2076(2003).			
DR AB016926; AAC0528.1;			
KW Complete proteome; Hypothetical protein.			
SQ SEQUENCE 419 AA; 48881 MW; D503382A1524C387 CRC64;			
Query Match 51.1%; Score 46%; DB 2; Length 419;			
Best Local Similarity 50.0%; Pred. No. 91;			
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;			

Qy 5 RQLAIGDDNIRK 18  
 Db 383 RKIVIGDDHRE 396

Search completed: November 10, 2004, 13:38:48  
 Job time : 23.9793 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw mode1  
 Run on: November 10, 2004, 12:32:37 : Search time 8.981. Seconds  
 (without alignments)  
 191.991 Million cell. updates/sec

Title: US-10-092-750-42  
 Perfect score: 123  
 Sequence: 1 GYSEAEQTEPLSTELGIASURSRVVA 26  
 Scoring table: BL0SUM62  
 Gapext 10.0 , Gapext 0.5  
 Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cggn2\_6/pododata/1/iaaa/5A.COMB.Pept:  
 2: /cggn\_6/pododata/1/iaaa/5B.COMB.Pept:  
 3: /cggn2\_6/pododata/1/iaaa/6A.COMB.Pept:  
 4: /cggn2\_6/pododata/1/iaaa/6B.COMB.Pept:  
 5: /cggn2\_6/pododata/1/iaaa/PPTUS.COMB.Pept:  
 6: /cggn2\_6/pododata/1/iaaa/backfiles.Pept:  
 \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	38.2	505	4 US-09-877-476-10	Sequence 10, App1
2	47	38.2	505	3 US-09-877-476-18	Sequence 18, App1
3	47	38.2	506	3 US-08-888-998-2	Sequence 2, App1
4	47	38.2	506	3 US-09-362-633-2	Sequence 2, App1
5	47	38.2	506	4 US-09-877-476-2	Sequence 2, App1
6	47	38.2	506	4 US-09-877-476-8	Sequence 8, App1
7	47	38.2	506	4 US-09-877-476-12	Sequence 12, App1
8	47	38.2	506	4 US-09-877-476-14	Sequence 14, App1
9	47	38.2	506	4 US-09-877-476-16	Sequence 16, App1
10	47	38.2	506	4 US-09-877-476-20	Sequence 20, App1
11	47	38.2	506	4 US-09-877-476-22	Sequence 22, App1
12	47	38.2	506	4 US-09-877-476-24	Sequence 24, App1
13	47	38.2	506	4 US-09-877-476-26	Sequence 26, App1
14	47	38.2	506	4 US-09-877-476-30	Sequence 30, App1
15	47	38.2	506	4 US-09-877-476-36	Sequence 36, App1
16	47	38.2	506	4 US-09-877-476-38	Sequence 38, App1
17	38.2	506	4 US-09-877-476-40	Sequence 40, App1	
18	47	38.2	740	1 US-08-309-512-10	Sequence 10, App1
19	47	38.2	740	5 PCT-US2-08756A-10	Sequence 10, App1
20	45	36.6	211	4 US-09-98-552A-302	Sequence 90z, App1
21	45	36.6	765	1 US-08-309-512-5	Sequence 5, App1
22	45	36.6	765	5 PCT-US2-08756A-5	Sequence 5, App1
23	44	35.8	885	4 US-09-248-796A-14427	Sequence 14,227, A
24	43	35.0	74	4 US-09-252-991A-22657	Sequence 22657, A
25	43	35.0	78	4 US-09-621-976-5620	Sequence 5620, App1
26	43	35.0	85	4 US-09-70-167-16286	Sequence 46,816, App1
27	35.0	113	4 US-09-886-319A-25	Sequence 25, App1	

#### ALIGNMENTS

RESULT 1	US-09-877-476-10	;	SEQUENCE 10, Application US/09877476
		;	GENERAL INFORMATION:
		;	PATENT NO. 6713664
		;	APPLICANT: Jaworski, Jan G.
		;	ATTORNEY: Blacklock, Brenda J.
		;	TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
		;	TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
		;	FILE REFERENCE: 07148-108001
		;	CURRENT APPLICATION NUMBER: US/09/877,476
		;	CURRENT FILING DATE: 2001-06-08
		;	PRIOR APPLICATION NUMBER: US 60/210,326
		;	PRIOR FILING DATE: 2000-06-08
		;	NUMBER OF SEQ ID NOS: 56
		;	SOFTWARE: FASTSEQ For Windows Version 4.0
		;	SEQ ID NO 10
		;	LENGTH: 505
		;	TYPE: PRT
		;	ORGANISM: Artificial Sequence
		;	FEATURE:
		;	OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3' 431 amino acids from B. napus
		;	OTHER INFORMATION: NO:2; and 3, 431 amino acids from B. napus
		;	OTHER INFORMATION: designated At74
		;	US-09-877-476-10
		;	Query Match 38.2%; Score 47; DB 4; Length 505;
		;	Best Local Similarity 71.4%; Pred. No. 19;
		;	Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy	9 FPISTFLIGIAsRL 22	Db	23 FPIIAFLAGKASRL 36
RESULT 2	US-09-877-476-18	;	SEQUENCE 18, Application US/09877476
		;	GENERAL INFORMATION:
		;	PATENT NO. 6713664
		;	APPLICANT: Jaworski, Jan G.
		;	ATTORNEY: Blacklock, Brenda J.
		;	TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
		;	TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
		;	FILE REFERENCE: 07148-108001
		;	CURRENT APPLICATION NUMBER: US/09/877,476
		;	CURRENT FILING DATE: 2001-06-08
		;	PRIOR APPLICATION NUMBER: US 60/210,326
		;	PRIOR FILING DATE: 2000-06-08
		;	NUMBER OF SEQ ID NOS: 56

SOFTWARE: FastSEQ for windows Version 4.0  
 / SEQ ID NO 18  
 / LENGTH: 505  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID  
 / OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus  
 / OTHER INFORMATION: elongate KCS (SEQ ID NO:4) having a mutation at  
 / OTHER INFORMATION: residue 306; designated At74 G306D; hypothetical  
 .US-09-877-476-18

RESULT 4  
 / Sequence 2, Application US/09362633  
 / Patent No. 6184355  
 / GENERAL INFORMATION:  
 / APPLICANT: JAMES, Douglas W.  
 / APPLICANT: LIM, Eda  
 / APPLICANT: KELLER, Janis  
 / APPLICANT: DOONER, Hugo K.  
 / TITLE OF INVENTION: FAE1 GENES AND THEIR USES  
 / NUMBER OF SEQUENCES: 8  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Townsend and Townsend Khourie and Crew  
 / STREET: Stewart Street Tower, One Market Plaza  
 / CITY: San Francisco  
 / STATE: California  
 / COUNTRY: US  
 / ZIP: 94105-1493  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC Compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.30  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/09/362,633  
 / FILING DATE:  
 / CLASSIFICATION:  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: 08/888,998  
 / FILING DATE:  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Bastian, Kevin L.  
 / REGISTRATION NUMBER: 34,774  
 / REFERENCE/DOCKET NUMBER: 12176-004300  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (415) 543-5043  
 / TELEFAX: (415) 543-5043  
 / INFORMATION FOR SEQ ID NO: 2:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 506 amino acids  
 / TYPE: amino acid  
 / STRANDEDNESS: not relevant  
 / TOPOLOGY: not relevant  
 / MOLECULE TYPE: peptide  
 / FEATURES:  
 / NAME/KEY: Protein  
 / LOCATION: 1..506  
 / OTHER INFORMATION: /note= "Amino acid sequence of FAE1  
 / OTHER INFORMATION: protein."  
 US-09-362-633-2

Query Match 38.2%; Score 47; DB 3; Length 506;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 5  
 US-09-877-476-2  
 / Sequence 2, Application US/09877476  
 / Patent No. 6713664  
 / GENERAL INFORMATION:

Software: FastSEQ for windows Version 4.0  
 / SEQ ID NO 18  
 / LENGTH: 505  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID  
 / OTHER INFORMATION: NO:2) and 3' 431 amino acids from B. napus  
 / OTHER INFORMATION: elongate KCS (SEQ ID NO:4) having a mutation at  
 / OTHER INFORMATION: residue 306; designated At74 G306D; hypothetical  
 .US-08-888-998-2

RESULT 3  
 / Sequence 2, Application US/08888998  
 / Patent No. 6124524  
 / GENERAL INFORMATION:  
 / APPLICANT: JAMES, Douglas W.  
 / APPLICANT: LIM, Eda  
 / APPLICANT: KELLER, Janis  
 / APPLICANT: DOONER, Hugo K.  
 / TITLE OF INVENTION: FAE1 GENES AND THEIR USES  
 / NUMBER OF SEQUENCES: 8  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Townsend and Townsend Khourie and Crew  
 / STREET: Stewart Street Tower, One Market Plaza  
 / CITY: San Francisco  
 / STATE: California  
 / COUNTRY: US  
 / ZIP: 94105-1493  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC Compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.30  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/888,998  
 / FILING DATE: 07-JUL-1997  
 / CLASSIFICATION: 800  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: US 08/329,603  
 / FILING DATE: 26-OCT-1994  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Bastian, Kevin L.  
 / REGISTRATION NUMBER: 34,774  
 / REFERENCE/DOCKET NUMBER: 12176-004300  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (415) 543-5043  
 / TELEFAX: (415) 543-5043  
 / INFORMATION FOR SEQ ID NO: 2:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 506 amino acids  
 / TYPE: amino acid  
 / STRANDEDNESS: not relevant  
 / TOPOLOGY: not relevant  
 / MOLECULE TYPE: peptide  
 / FEATURES:  
 / NAME/KEY: Protein  
 / LOCATION: 1..506  
 / OTHER INFORMATION: /note= "Amino acid sequence of FAE1  
 / OTHER INFORMATION: protein."  
 US-09-362-633-2

Query Match 38.2%; Score 47; DB 3; Length 506;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 5  
 US-09-877-476-2  
 / Sequence 2, Application US/09877476  
 / Patent No. 6713664  
 / GENERAL INFORMATION:

Query Match 38.2%; Score 47; DB 3; Length 506;

APPLICANT: Jaworski, Jan G  
 APPLICANT: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 2  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 US-09-877-476-2

Query Match 38.2%; Score: 47; DB: 4; Length: 506;  
 Best Local Similarity: 71.4%; Pred. No.: 19; Indels: 3; Mismatches: 1; Gaps: 0;  
 Matches: 10; Conservative: 1; Gaps: 0;  
 Qy 9 FPLSTFLGIGASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 6  
 US-09-877-476-8  
 Sequence 8, Application US/09877476  
 ; Patent No. 6713664  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaworski, Jan G.  
 ; APPLICANT: Blacklock, Brenda J.  
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
 ; FILE REFERENCE: 07148-108001  
 ; CURRENT APPLICATION NUMBER: US/09/877,476  
 ; CURRENT FILING DATE: 2001-06-08  
 ; PRIOR APPLICATION NUMBER: US 60/210,326  
 ; PRIOR FILING DATE: 2000-06-08  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 8  
 ; LENGTH: 506  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (0) ... (0)  
 ; OTHER INFORMATION: Xaa = Ala or Thr  
 ; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID  
 ; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus  
 ; OTHER INFORMATION: (NO:2) and 3' 392 amino acids from B. napus  
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At114  
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At114  
 ; SEQ ID NO: 8

Query Match 38.2%; Score: 47; DB: 4; Length: 506;  
 Best Local Similarity: 71.4%; Pred. No.: 19; Indels: 3; Mismatches: 1; Gaps: 0;  
 Matches: 10; Conservative: 1; Gaps: 0;  
 Qy 9 FPLSTFLGIGASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 7  
 US-09-877-476-12  
 Sequence 12, Application US/09877476  
 ; Patent No. 6713664  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaworski, Jan G.  
 ; APPLICANT: Blacklock, Brenda J.  
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
 ; FILE REFERENCE: 07148-108001

APPLICANT: Jaworski, Jan G  
 APPLICANT: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 TITLE OF INVENTION: SYNTHASE POLYPEPTIDES

FILE REFERENCE: 07148-108001  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID  
 OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus  
 OTHER INFORMATION: elongase KCS (SEQ ID NO:4); having mutations at  
 OTHER INFORMATION: residues 91 and 92; designated At114 L91C K92R  
 US-09-877-476-12

Query Match 38.2%; Score: 47; DB: 4; Length: 506;  
 Best Local Similarity: 71.4%; Pred. No.: 19; Indels: 3; Mismatches: 1; Gaps: 0;  
 Matches: 10; Conservative: 1; Gaps: 0;  
 Qy 9 FPLSTFLGIGASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 8  
 US-09-877-476-14  
 Sequence 14, Application US/09877476  
 ; Patent No. 6713664  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaworski, Jan G.  
 ; APPLICANT: Blacklock, Brenda J.  
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
 ; FILE REFERENCE: 07148-108001  
 ; CURRENT APPLICATION NUMBER: US/09/877,476  
 ; CURRENT FILING DATE: 2001-06-08  
 ; PRIOR APPLICATION NUMBER: US 60/210,326  
 ; PRIOR FILING DATE: 2000-06-08  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 14  
 ; LENGTH: 506  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID  
 ; OTHER INFORMATION: NO:2) and 3' 392 amino acids from B. napus  
 ; OTHER INFORMATION: elongase KCS (SEQ ID NO:4); having a mutation at  
 ; OTHER INFORMATION: position 92; designated At114 K92R  
 US-09-877-476-14

Query Match 38.2%; Score: 47; DB: 4; Length: 506;  
 Best Local Similarity: 71.4%; Pred. No.: 19; Indels: 3; Mismatches: 1; Gaps: 0;  
 Matches: 10; Conservative: 1; Gaps: 0;  
 Qy 9 FPLSTFLGIGASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 9  
 US-09-877-476-16  
 Sequence 16, Application US/09877476  
 ; Patent No. 6713664  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jaworski, Jan G.  
 ; APPLICANT: Blacklock, Brenda J.  
 ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
 ; FILE REFERENCE: 07148-108001

```

CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,326
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 16
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
residue 307; designated At114 G307D; hypothetical
US-09-877-476-16

Query Match Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;
Qy 9 FPLSTFLGIGASRL 22
Db 23 FPLTAFLAGKASRL 36

RESULT 10
US-09-877-476-20
Sequence 20, Application US/09877476
Patent No. 6713664
GENERAL INFORMATION:
APPLICANT: Jaworski, Jan G.
ATTORNEY OR AGENT FOR APPLICANT: Blacklock, Brenda J.
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
FILE REFERENCE: 07148-108001
CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,326
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 20
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
residue 91, 92 and 307; designated At114 L91C
OTHER INFORMATION: R92R G307D; hypothetical
US-09-877-476-20

Query Match Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;
Qy 9 FPLSTFLGIGASRL 22
Db 23 FPLTAFLAGKASRL 36

RESULT 11
US-09-877-476-22
Sequence 22, Application US/09877476
Patent No. 6713664
GENERAL INFORMATION:
APPLICANT: Jaworski, Jan G.
ATTORNEY OR AGENT FOR APPLICANT: Blacklock, Brenda J.
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
FILE REFERENCE: 07148-108001
CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 22
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
residue 91, 92 and 307; designated At114 K92R G307D;
OTHER INFORMATION: positions 92 and 307; designated At114
OTHER INFORMATION: hypothetical
US-09-877-476-22

Query Match Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;
Qy 9 FPLSTFLGIGASRL 22
Db 23 FPLTAFLAGKASRL 36

RESULT 12
US-09-877-476-24
Sequence 24, Application US/09877476
Patent No. 6713664
GENERAL INFORMATION:
APPLICANT: Jaworski, Jan G.
ATTORNEY OR AGENT FOR APPLICANT: Blacklock, Brenda J.
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
FILE REFERENCE: 07148-108001
CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,326
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 24
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At254
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At254
US-09-877-476-24

Query Match Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;
Qy 9 FPLSTFLGIGASRL 22
Db 23 FPLTAFLAGKASRL 36

RESULT 13
US-09-877-476-26
Sequence 26, Application US/09877476
Patent No. 6713664
GENERAL INFORMATION:
APPLICANT: Jaworski, Jan G.
ATTORNEY OR AGENT FOR APPLICANT: Blacklock, Brenda J.
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
FILE REFERENCE: 07148-108001
CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 26
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At254
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At254
US-09-877-476-26

Query Match Score 47; DB 4; Length 506;
Best Local Similarity 71.4%; Pred. No. 19;
Matches 10; Conservative 1; Mismatches 0; Gaps 0;
Qy 9 FPLSTFLGIGASRL 22
Db 23 FPLTAFLAGKASRL 36

```

RESULT 14  
US-09-877-476-30  
Sequence 30, Application US/09877476  
Patent No. 6713664  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Jan G.  
APPLICANT: Blacklock, Brenda J.  
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
FILE REFERENCE: 07148-108001  
CURRENT APPLICATION NUMBER: US/09/877,476  
CURRENT FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/210,326  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 56  
SEQ ID NO: 26  
LENGTH: 506  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: 5' 173 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3' 333 amino acids from B. napus At173  
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At173  
NAME/KEY: VARIANT  
LOCATION: (0) ... (0)  
OTHER INFORMATION: Xaa = Pro or Gln  
US-09-877-476-36

Query Match 38.2%; Score 47; DB 4; Length 506;  
Best Local Similarity 71.4%; Pred. No. 19;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 9 FPLSTFLIGIASRL 22  
Db 23 FPLTAFLAGKASRL 36

Search completed: November 10, 2004, 13:44:11  
Job time : 8.981 secs

---

RESULT 14-476-30  
Sequence 30, Application US/09877476  
Patent No. 6713664  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Jan G.  
APPLICANT: Blacklock, Brenda J.  
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
FILE REFERENCE: 07148-108001  
CURRENT APPLICATION NUMBER: US/09/877,476  
CURRENT FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/210,326  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 56  
SEQ ID NO: 30  
LENGTH: 506  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: 5' 399 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3' 107 amino acids from B. napus At399  
OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At399  
NAME/KEY: VARIANT  
LOCATION: (0) ... (0)  
OTHER INFORMATION: Xaa = Pro or Gln  
US-09-877-476-30

Query Match 38.2%; Score 47; DB 4; Length 506;  
Best Local Similarity 71.4%; Pred. No. 19;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 9 FPLSTFLIGIASRL 22  
Db 23 FPLTAFLAGKASRL 36

RESULT 15  
US-09-877-476-36  
Sequence 36, Application US/09877476  
Patent No. 6713664  
GENERAL INFORMATION:  
APPLICANT: Jaworski, Jan G.  
APPLICANT: Blacklock, Brenda J.  
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 16:36:12 ; Search time 27.9309 Seconds  
(without alignments)  
328.807 Million cell updates/sec

Title: US-10-092-750-42  
Perfect score: 123  
Sequence: 1 GVSEAGTFPLSTFLIGIASRLRSVA 26

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35325886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
 1: /cgcn2\_6/ptodata/1/pubpaas/pctus\_pubcomb.pep/\*  
 2: /cgcn2\_6/prodata/1/pubpaas/pctus\_new\_pub.pep/\*  
 3: /cgcn2\_6/prodata/1/pubpaas/us06\_pub.pep/\*  
 4: /cgcn2\_6/prodata/1/pubpaas/us05\_pubcomb.pep/\*  
 5: /cgcn2\_6/prodata/1/pubpaas/us07\_pub.pep/\*  
 6: /cgcn2\_6/prodata/1/pubpaas/pctus\_pubcomb.pep/\*  
 7: /cgcn2\_6/prodata/1/pubpaas/us08\_pub.pep/\*  
 8: /cgcn2\_6/prodata/1/pubpaas/us08\_pubcomb.pep/\*  
 9: /cgcn2\_6/prodata/1/pubpaas/us09\_pubcomb.pep/\*  
 10: /cgcn2\_6/prodata/1/pubpaas/us09\_pubcomb.pep/\*  
 11: /cgcn2\_6/prodata/1/pubpaas/us09c\_pubcomb.pep/\*  
 12: /cgcn2\_6/prodata/1/pubpaas/us09\_new\_pub.pep/\*  
 13: /cgcn2\_6/prodata/1/pubpaas/us09\_pubcomb.pep/\*  
 14: /cgcn2\_6/prodata/1/pubpaas/us10b\_pubcomb.pep/\*  
 15: /cgcn2\_6/prodata/1/pubpaas/us10c\_pubcomb.pep/\*  
 16: /cgcn2\_6/prodata/1/pubpaas/us10d\_pubcomb.pep/\*  
 17: /cgcn2\_6/prodata/1/pubpaas/us10\_new\_pub.pep/\*  
 18: /cgcn2\_6/prodata/1/pubpaas/us11\_new\_pub.pep/\*  
 19: /cgcn2\_6/prodata/1/pubpaas/us11\_new\_pub.pep/\*  
 20: /cgcn2\_6/prodata/1/pubpaas/us10\_pubcomb.pep/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	100.0	26	14 US-10-092-750-42	Sequence 42, Appl
2	51	41.5	87	17 US-10-425-115-164584	Sequence 264584,
3	50	40.7	264	15 US-10-424-599-155062	Sequence 155062,
4	48	39.0	894	16 US-10-437-963-119916	Sequence 119916,
5	47	38.2	505	9 US-09-877-476-10	Sequence 10, Appl
6	47	38.2	505	9 US-09-877-476-18	Sequence 18, Appl
7	47	38.2	505	16 US-10-758-524-10	Sequence 10, Appl
8	47	38.2	505	16 US-10-758-524-18	Sequence 18, Appl
9	47	38.2	506	9 US-09-877-476-2	Sequence 2, Appl
10	47	38.2	506	9 US-09-877-476-24	Sequence 8, Appl
11	47	38.2	506	9 US-09-877-476-12	Sequence 12, Appl
12	47	38.2	506	9 US-09-877-476-14	Sequence 14, Appl
13	47	38.2	506	9 US-09-877-476-16	Sequence 16, Appl

RESULT 1  
US-10-092-750-42

; Sequence 42, Application US-10092750  
; Publication No. US2003032157A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; ATTORNEY: Alpin, Julia  
; ATTORNEY: Wright, Martin C.  
; TITLE OF INVENTION: Peptidides Interactive with BCL-X1  
; FILE REFERENCE: 50036/50002  
; CURRENT APPLICATION NUMBER: US-10-092-750  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US-10-092-750  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-42

## ALIGNMENTS

Query Match Local Similarity 100.0%; Score 123; DB 14; Length 26;  
Best Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVSBEAGTFPLSTFLIGIASRLRSVA 26  
Db 1 GVSBEAGTFPLSTFLIGIASRLRSVA 26

RESULT 2  
US-10-425-115-264584  
; Sequence 264584, Application US-10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yinhua  
 APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants

FILE REFERENCE: 38-21(53221) B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 SEQ ID NO: 119916  
 LENGTH: 894  
 TYPE: PRT  
 ORGANISM: Oryza sativa

NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO: 261584  
 LENGTH: 87

TYPE: PRT  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_23086C.1.pep

FEATURE: Zea mays

OTHER INFORMATION: Clone ID: MRT4577\_172912C.1.pep  
 US-10-425-115-264584

Query Match Score 51; DB 17; Length 87;  
 Best Local Similarity 44.8%; Pred. No. 3;  
 Matches 13; Conservative 2; Mismatches 4; Indels 10; Gaps 1;

Db 50 AEGTFPLSTF-----LIGIASRL 23  
 Qy 5 AEGTFPLSTF-----LIGIASRL 23

RESULT 3  
 US-10-424-599-155062  
 Sequence 155052, Application US/10424599  
 Publication No. US20040031072A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yinhua  
 APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Plants

FILE REFERENCE: 38-21(53223) B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 SEQ ID NO: 285684  
 LENGTH: 264  
 TYPE: PRT  
 FEATURE: Glycine max

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_111042C.1.pep  
 US-10-424-599-155062

Query Match Score 50; DB 15; Length 264;  
 Best Local Similarity 62.5%; Pred. No. 16;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 21 TFPPLSTFLIGIASRL 23  
 Qy 21 TIPPLTLIGIVSRIR 36

RESULT 4  
 US-10-437-963-119916  
 Sequence 119916, Application US/10437963  
 Publication No. US2005012334A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yinhua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

; SEQ ID NO: 18 ; NUMBER OF SEQ ID NOS: 56 ; SOFTWARE: FastSEQ for Windows Version 4.0
; LENGTH: 505 ; SEQ ID NO: 18
; TYPE: PRT ; LENGTH: 505
; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3', 431 amino acids from B. napus OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having a mutation at residue 306; designated Ac74 G306D; hypothetical US-03-877-476-18 ; OTHER INFORMATION: residue 306; designated At74 G306D; hypothetical US-10-758-524-18

Query Match 38.2%; Score 47; DB 9; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 10; Conservative 1; Gaps 0;

Qy 9 FPLSTFLGGASRL 22 Query Match 38.2%; Score 47; DB 16; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 10; Conservative 1; Gaps 0;

Db 23 FPLTAFLAGGASRL 36 Qy 9 FPLSTFLGGASRL 22
Db 23 FPLTAFLAGGASRL 36

RESULT 7 Query Match 38.2%; Score 47; DB 16; Length 505;
US-10-758-524-10 Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
; Sequence 10, Application US/10758524 ; SEQ ID NO: 2 ; Application US/09877476
; Publication No. US20040139498A1 ; Patent No. US0020049994A1
; GENERAL INFORMATION: ; GENERAL INFORMATION: ; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J. ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES ; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES
; FILE REFERENCE: 07148-108002 ; FILE REFERENCE: 07148-108001
; CURRENT APPLICATION NUMBER: US/10/758,524 ; CURRENT APPLICATION NUMBER: US/09/877,476
; PRIOR APPLICATION NUMBER: 2001-06-08 ; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/210,326 ; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08 ; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56 ; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0 ; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 10 ; SEQ ID NO: 2
; LENGTH: 505 ; LENGTH: 506
; TYPE: PRT ; TYPE: PRT
; ORGANISM: Artificial Sequence ; ORGANISM: Arabidopsis thaliana
; FEATURE: ; OTHER INFORMATION: 5' 74 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3', 431 amino acids from B. napus OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At74 US-10-758-524-10 ; OTHER INFORMATION: residue 306; designated Ac74 G306D; hypothetical US-10-758-524-18

Query Match 38.2%; Score 47; DB 16; Length 505;
Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 10; Conservative 1; Gaps 0;

Qy 9 FPLSTFLGGASRL 22 Query Match 38.2%; Score 47; DB 9; Length 506;
Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
Matches 10; Conservative 1; Gaps 0;

Db 23 FPLTAFLAGGASRL 36 Qy 9 FPLSTFLGGASRL 22
Db 23 FPLTAFLAGGASRL 36

RESULT 10 Query Match 38.2%; Score 47; DB 9; Length 506;
US-09-877-476-8 Best Local Similarity 71.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;
; Sequence 8, Application US/09877476 ; SEQ ID NO: 8 ; Application US/09877476
; Publication No. US0020049994A1 ; Patent No. US0020049994A1
; GENERAL INFORMATION: ; GENERAL INFORMATION: ; APPLICANT: Jaworski, Jan G.
; APPLICANT: Blacklock, Brenda J. ; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES ; FILE REFERENCE: 07148-108001
; FILE REFERENCE: 07148-108002 ; CURRENT APPLICATION NUMBER: US/09/877,476
; CURRENT FILING DATE: 2001-06-08 ; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 60/210,326 ; PRIOR APPLICATION NUMBER: US 60/210,326
; PRIOR FILING DATE: 2000-06-08 ; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 56 ; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0 ; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 8 ; LENGTH: 506
; TYPE: PRT ; TYPE: PRT
; ORGANISM: Artificial Sequence

```

FEATURE:  
 / NAME/KEY: VARIANT  
 / LOCATION: (0)...(0)  
 OTHER INFORMATION: Xaa = Ala or Thr  
 OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID NO:1);  
 OTHER INFORMATION: 3' 392 amino acids from B. napus (SEQ ID NO:2);  
 OTHER INFORMATION: 5' 114 amino acids from B. napus (SEQ ID NO:3);  
 OTHER INFORMATION: elongase KCS (SEQ ID NO:4); designated At114 US-09-877-476-8

Query Match 38.2%; Score 47; DB 9; Length 506;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 9 FPLSTFLGLASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 11  
 US-09-877-476-12  
 Sequence 12, Application US/09877476  
 Patent No. US2002004994A1  
 GENERAL INFORMATION:  
 APPLICANT: Jaworski, Jan G.  
 ATTORNEY: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 SYNTHASE POLYPEPTIDES  
 FILE REFERENCE: 07148-108001  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 12  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID NO:1);  
 OTHER INFORMATION: 3' 392 amino acids from B. napus (SEQ ID NO:2);  
 OTHER INFORMATION: elongase KCS (SEQ ID NO:4); having mutations at residues 91 and 92; designated At114 L91C K92R US-09-877-476-12

Query Match 38.2%; Score 47; DB 9; Length 506;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 9 FPLSTFLGLASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 12  
 US-09-877-476-14  
 Sequence 14, Application US/09877476  
 Patent No. US2002004994A1  
 GENERAL INFORMATION:  
 APPLICANT: Jaworski, Jan G.  
 ATTORNEY: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 SYNTHASE POLYPEPTIDES  
 FILE REFERENCE: 07148-108001  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 14  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Artificial Sequence

RESULT 13  
 US-09-877-476-16  
 Sequence 16, Application US/09877476  
 Patent No. US2002004994A1  
 GENERAL INFORMATION:  
 APPLICANT: Jaworski, Jan G.  
 ATTORNEY: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 SYNTHASE POLYPEPTIDES  
 FILE REFERENCE: 07148-108001  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 16  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID NO:1);  
 OTHER INFORMATION: 3' 392 amino acids from B. napus (SEQ ID NO:2);  
 OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutation at residue 307; designated At114 G307D; hypothetical OTHER INFORMATION: residue 307; designated At114 G307D; hypothetical US-09-877-476-16

Query Match 38.2%; Score 47; DB 9; Length 506;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 9 FPLSTFLGLASRL 22  
 Db 23 FPLTAFLAGKASRL 36

RESULT 14  
 US-09-877-476-20  
 Sequence 20, Application US/09877476  
 Patent No. US2002004994A1  
 GENERAL INFORMATION:  
 APPLICANT: Jaworski, Jan G.  
 ATTORNEY: Blacklock, Brenda J.  
 TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
 SYNTHASE POLYPEPTIDES  
 FILE REFERENCE: 07148-108001  
 CURRENT APPLICATION NUMBER: US/09/877,476  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08  
 NUMBER OF SEQ ID NOS: 56  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 20  
 LENGTH: 506  
 TYPE: PRT  
 ORGANISM: Artificial Sequence

```

FEATURE
; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
; NO:2) and 3' 392 amino acids from B. napus
; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
; positions 9, 92 and 307; designated At14 L91C
; OTHER INFORMATION: K92R G307D; hypothetical
US-09-877-476-20

```

```

Query Match      38.2%;  Score 47;  DB 9;  Length 506;
Best Local Similarity 71.4%;  Pred. No. 1e+02;  Gaps 0;
Matches 10;  Conservative 1;  Mismatches 3;  Indels 0;  Gaps 0;
Qy   9 FPLSTFLLGASRL 22
      |||:||| | | | |
Db    23 FPLTAFLAGKASRL 36

```

## RESULT 15

```

US-09-877-476-22
Sequence 22, Application US/09877476
Parent No. US2002004994A1
GENERAL INFORMATION:
APPLICANT: Jaworski, Jan G.
APPLICANT: Blacklock, Brenda J.
TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA
SYNTHASE POLYPEPTIDES
FILE REFERENCE: 07148-108001
CURRENT APPLICATION NUMBER: US/09/877,476
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 60/210,326
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 22
LENGTH: 506
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID
NO:2) and 3' 392 amino acids from B. napus
OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at
positions 9, 92 and 307; designated At14 L91C
OTHER INFORMATION: K92R G307D; hypothetical
US-09-877-476-22

```

```

Query Match      38.2%;  Score 47;  DB 9;  Length 506;
Best Local Similarity 71.4%;  Pred. No. 1e+02;  Gaps 0;
Matches 10;  Conservative 1;  Mismatches 3;  Indels 0;  Gaps 0;
Qy   9 FPLSTFLLGASRL 22
      |||:||| | | | |
Db    23 FPLTAFLAGKASRL 36

```

Search completed: November 11, 2004, 01:28:23  
Job time : 27.9809 secs

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 12:29:32 ; Search time 5.79275 Seconds

(without alignments) 431.857 Million cell updates/sec

Title: US-10-092-750-42

Perfect score: 123

Sequence: 1 GYSEAEGTEPLSTELLGIASRLRSVA 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:

283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	50.5	41.1	209	2	B90497		conserved hypothetical protein [imported] - <i>Sulfolobus solfataricus</i>
2	48	39.0	154	2	H72269		probable membrane protein
3	47.5	38.6	1467	2	T48162		hypothetical protein
4	47	38.2	275	2	A56641		probable membrane protein
5	47	38.2	487	2	G83827		stage V sporulation
6	47	38.2	506	2	T05272		fatty acid elongase
7	47	38.2	513	2	T43434		translation elongase
8	47	38.2	1047	2	T41343		probable translase
9	46.5	37.8	153	2	AB1414		hypothetical protein
10	46.5	37.8	478	2	T45661		hypothetical protein
11	46.5	37.8	839	2	AC2637		large atp-dependent
12	46.5	37.8	848	2	B97419		probable ATP-dependent
13	46	37.4	113	2	S59116		programmed cell death
14	46	37.4	126	2	AB0326		membrane protein
15	46	37.4	254	2	G70386		flagellar motor protein
16	46	37.4	358	2	H90052		conserved hypothetical protein
17	46	37.4	376	2	G71290		flagellar biosynthesis
18	46	37.4	485	2	G85943		probable transport
19	46	37.4	485	2	B65072		probable transport
20	46	37.4	495	2	C91098		aldehyde-ferredoxin
21	46	37.4	581	2	C71078		probable ABC transporter
22	46	37.4	1469	2	H96622		myb-like DNA-binding
23	45.5	37.0	556	2	T38479		probable outer membrane
24	45.5	37.0	853	2	A71339		apoptotic cell death
25	45	36.6	113	2	B54437		integral membrane protein
26	45	36.6	123	2	AC3331		endonuclease III
27	45	36.6	209	2	D72029		[hypothetical protein]
28	45	36.6	209	2	C86595		conserved hypothetical protein MJ0281
29	45	36.6	466	2	T264558		Query Match

#### ALIGNMENTS

RESULT 1	Query	Match	Length	DB	Score	DB 2	Length	DB 3	Score	DB 4	Length
B90497		conserved hypothetical protein [imported] - <i>Sulfolobus solfataricus</i>		C:Species: <i>Sulfolobus solfataricus</i>		C:Accession: B90497		C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004		C:Residues: 1-209 <KUR>	
R1she, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awaye, M.J.; Chan-Jong, I.; Jeffries, A.C.; Koza, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.; Garrett, R.A.; Regan, M.A.; Sersen, C.W.; Van der Oost, J.		A:Residues: 1-209		A:Cross-references: UNIPROT:Q97U88 ; GB:AE006641 ; NID:913816556 ; PID:AAK43233.1 ; GSPDB:GI		A:Description: <i>Sulfolobus solfataricus</i> complete genome		A:Gene: SS03132		A:Accession: B90497	
C:Status: preliminary		A:Molecule type: DNA		C:Generics:		C:Superfamily: glycosidase		C:Accession: B90497		C:Residues: 1-209	
Qy		1 GYSEAESTEPPLSTEL---		Qy		1 GYSEAESTEPPLSTEL---		Qy		1 GYSEAESTEPPLSTEL---	
Db		62 GYNQIIRDFPSSTFLNEKDGLLKLASSMA		Db		62 GYNQIIRDFPSSTFLNEKDGLLKLASSMA		Db		92	
RESULT 2	Query	Match	Length	DB	Score	DB 2	Length	DB 3	Score	DB 4	Length
H72269		hypothetical protein APE147 - <i>Aeropyrum pernix</i> (strain K1)		C:Species: <i>Aeropyrum pernix</i>		C:Accession: H72269		C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004		C:Residues: 1-154 <KRW>	
R:Kawarabayashi, Y.; Hino, Y.; Horioka, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funaiishi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kikuchi, A.; Takamiya, M.; Masuda, S.; Funaiishi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kikuchi, A.; Garrett, R.A.; Regan, M.A.; Sersen, C.W.; Van der Oost, J.		A:Residues: 1-154		A:Cross-references: UNIPROT:Q9YFV3 ; DDBJ:AP000058 ; PID:NID:95103388 ; PMID:10382966		A:Description: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, <i>Aeropyrum pernix</i> (strain K1)		A:Gene: AP0147		A:Accession: H72269	
C:Status: preliminary		A:Molecule type: DNA		C:Generics:		C:Superfamily: conserved hypothetical protein MJ0281		C:Accession: H72269		C:Residues: 1-154	
Qy		1 GYSEAESTEPPLSTEL---		Qy		1 GYSEAESTEPPLSTEL---		Qy		1 GYSEAESTEPPLSTEL---	
Db		62 GYNQIIRDFPSSTFLNEKDGLLKLASSMA		Db		92		Db		Query Match	

Best Local Similarity 42.3%; Pred. No. 3.9; Mismatches 11; Indels 0; Gaps 0;

Matches 11; Conservative 4; Amino acid type: DNA

A;Title: Complete Genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its plasmid pBAM132

A;Reference number: AB3650; PMID:20512582; MUID:11058132

A;Accession: G33827

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-87 <STO>

A;Cross-references: UNIPROT:Q9KCZ6; GB:AP001512; PID:BA000004; NID:910174030; PIDN:BA0514

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1423

RESULT 3

T48162 hypothetical protein T10008.110 - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: T48162

R;Bevan, M.; Pohl, T.; Weizenegger, T.; Barcroft, I.; Meves, H.W.; Lemcke, K.; Mayer, K.

Submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24486

A;Accession: T48162

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-1467 <BEV>

A;Cross-references: UNIPROT:Q9M033; EMBL:AU161746

A;Experimental source: cultivar Columbia; BAC clone T1008

C;Genetics:

A;Map position: 5

A;Introns: 68%; 121/3; 152/3; 178/1; 234/2; 291/3; 331/3; 516/3; 698/3; 721/3; 760/3; 82

A;Note: T1008.110

Query Match Score 38.6%; Best Local Similarity 58.3%; Pred. No. 48; Matches 14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Query Match Score 47.5%; Best Local Similarity 58.3%; Pred. No. 48; Matches 14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Query Match Score 47.5%; Best Local Similarity 58.3%; Pred. No. 48; Matches 14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Db 1146 V1QAGAFAISDFILEILSRVLS 1169

RESULT 4

A56641 probable membrane transport protein - *Clostridium perfringens*

C;Species: *Clostridium perfringens*

C;Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 09-Jul-2004

C;Accession: A56641

R;Holck, A.L.; Blom, H.

DNA Seq. 3, 191-194, 1992

A;Title: The nucleotide sequence of a putative membrane transport gene from *Clostridium* A;Reference number: A56641; PMID:9313001; MUID:1472712

A;Accession: A56641

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-275 <HOL>

A;Cross-references: UNIPROT:Q06111; GB:X66092; NID:g296355; PIDN:CAA446387.1; PID:g293556

A;Note: nucleotide sequence not given; conceptual translation not complete

C;Superfamily: maltose transport protein malaG

Query Match Score 38.2%; Best Local Similarity 38.5%; Pred. No. 10; Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Query Match Score 47%; Best Local Similarity 38.5%; Pred. No. 10; Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Query Match Score 47%; Best Local Similarity 38.5%; Pred. No. 10; Matches 10; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Db 146 GIVSAFGTFPLSTFLGIASRLRSVA 171

RESULT 5

G83827 stage V sporulation protein AF BH1423 [imported] - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004

C;Accession: G33827

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fuji, N.; Hirai, Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete Genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its plasmid pBAM132

A;Reference number: AB3650; PMID:20512582; MUID:11058132

A;Accession: G33827

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-87 <STO>

A;Cross-references: UNIPROT:Q9KCZ6; GB:AP001512; PID:BA000004; NID:910174030; PIDN:BA0514

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1423

RESULT 6

T05272 fatty acid elongase 1 - *Arabidopsis thaliana*

N;Alternative names: Protein T4120.100

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T05272

R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clercq, R.; Deewes, H.W.; Mayer, K.F.X.; Schueler, C.

Submitted to the Protein Sequence Database, September 1998

A;Reference number: Z15406

A;Accession: T05272

A;Molecule type: DNA

A;Residues: 1-506 <BEV>

A;Cross-references: UNIPROT:Q38860; EMBL:AL023094

A;Experimental source: cultivar Columbia; BAC clone T4L20

C;Genetics:

A;Map position: 4

A;Note: T4120.100

C;Superfamily: very-long-chain 3-ketoacyl-CoA synthase

Query Match Score 38.2%; Best Local Similarity 71.4%; Pred. No. 19; Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match Score 38.2%; Best Local Similarity 71.4%; Pred. No. 19; Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match Score 38.2%; Best Local Similarity 71.4%; Pred. No. 19; Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match Score 38.2%; Best Local Similarity 71.4%; Pred. No. 19; Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 9 FPLSTFLGIASRL 22

Db 23 FPLTAFLLAGSRL 36

RESULT 7

T34343 translation elongation factor eEF-3 - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C;Accession: T43434

R;Uritani, M.; Shoumura, Y.; Yamada, S.

Biosci. Biotechnol. Biochem. 63, 769-772, 1999

A;Title: Detection and analysis of translation elongation factor 3 genes from various yeasts

A;Reference number: Z22510; MUID:93290044; PMID:10361693

A;Accession: T43434

A;Status: Preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: 1-513 <URI>

A;Cross-references: UNIPROT:OP4489; EMBL:AB018538; PID:93776153; PIDN:BA33896.1; PID:G33827

C;Genetics:

A;Gene: EF\_3

C;Superfamily: translation elongation factor 3; ATP-binding cassette homology

Query Match Score 38.2%; Best Local Similarity 40.0%; Pred. No. 20; Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa  
submitted to the Protein Sequence Database, November 1999  
A;Reference number: 223010  
A;Accession: T45661  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q9SD48; EMBL:AL133292  
C;Genetics:  
A;Map position: 3  
A;Introns: 255/3  
A;Note: F13I12.300

RESULT 8  
T41343 probable translation elongation factor EF-3 - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Accession: T41343  
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Voilcaert, G.  
A;Reference number: Z21970  
A;Accession: T41343  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1047 <LYN>  
A;Cross-references: UNIPROT:Q944B9; EMBL:AL035076; PIDN:CAA22654.1; GSPDB:GN000666;  
A;Experimental source: strain 972h.; cosmid c417  
C;Genetics:  
A;Gene: SPPB:SPEC417.08  
A;Map position: 1  
C;Superfamily: translation elongation factor 3; ATP-binding cassette homology  
Query Match Score 47; DB 2; Length 1047;  
Best Local Similarity 40.0%; Pred. No. 41; Gaps 0;  
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 GVSEAGTFPLSTFLGIASRLRSV 25  
Db 648 GASEMEEFKPBPGLFLEGVTKQRAl 672

RESULT 9  
AB1415 hypothetical protein lmo2723 [imported] - Listeria monocytogenes (strain EGDe-e)  
C;Species: Listeria monocytogenes  
C;Accession: AB1415  
R;Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher  
D.; Dominguez-Bernal, G.; Duchaoud, E.; Durand, L.; Dussurgeot, O.; Entian, K.D.; Feihl, H.  
Science 294, 849-852, 2001  
A;Authors: Kruhn, M.; Kunst, F.; Kuraphat, G.; Madueno, B.; Maitouram, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AB1415  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:QBY3VB; GB:NC\_003210; PIDN:CAD00936.1; PID:916412223; GSPDB:  
A;Experimental source: strain EGDe-e  
A;Genes: lmo2723

Query Match Score 37.8%; DB 2; Length 153;  
Best Local Similarity 35.7%; Pred. No. 6.7%; Gaps 2;  
Matches 15; Conservative 3; Mismatches 5; Indels 19; Gaps 2;

Qy 4 BAEGTFPLSTF---LLGIAS-----RLRSVA 26  
Db 34 DMEGTHLGAEKDVILGVASFYPEKSTVMPAOYRIGVVA 75

RESULT 10  
T45661 hypothetical protein F13I12.300 - Arabidopsis thaliana (mouse-ear cress)  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Accession: T45661  
A;Cross-references: UNIPROT:Q8U106; GB:AE007869; PIDN:AAK86307.1; PMID:915155423; GSPDB:G

R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa  
submitted to the Protein Sequence Database, November 1999  
A;Reference number: 223010  
A;Accession: T45661  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q9SD48; EMBL:AL133292  
C;Genetics:  
A;Map position: 3  
A;Introns: 255/3  
A;Note: F13I12.300

Query Match Score 37.8%; DB 2; Length 478;  
Best Local Similarity 50.0%; Pred. No. 22; Gaps 1;  
Matches 15; Conservative 1; Mismatches 9; Indels 5; Gaps 1;

Qy 2 VSEAGTFPL---STFLGIASRSVA 26  
Db 285 VSSSESTFPPLSAKURLQGIKFRURSDA 314

RESULT 11  
AC2637 large atp-dependent helicase-related protein [imported] - Agrobacterium tumefaciens (strain C5t)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AC2637  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Nonks, D.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
Science 294, 2317-2333, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Gordon-Kamm, I.  
ster, E. W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AC2637  
A;Status: preliminary  
A;Molecule type: DNA  
A;Cross-references: UNIPROT:Q8U106; GB:AE008688; PIDN:AAI41513.1; PID:917738842; GSPDB:G  
C;Genetics:  
A;Map position: circular chromosome

Query Match Score 37.8%; DB 2; Length 839;  
Best Local Similarity 52.2%; Pred. No. 39; Gaps 4;  
Matches 12; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

Qy 3 SEAGTFPLSTFLGIASRSV 25  
Db 570 SYAGGKFPLSTVLI---ADQVRHSM 589

RESULT 12  
B97419 probable ATP-dependent helicase mjo294 [imported] - Agrobacterium tumefaciens (strain C5t)  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: B97419  
R;Goodier, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, C.; Mazkelz, B.;  
A; Liu, F.; Wollam, C.; Dougherty, D.; Scott, C.; Lappas, C.; Scott, C.; Lappas, C.; Mazkelz, B.;  
Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: B97419  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-848 <CUR>  
C;Genetics:

A;Gene: AGR C\_874  
A;Map Position: circular chromosome

Query Match Score 46.5; DB 2; Length 848;  
Best Local Similarity 52.2%; Pred. No. 39;  
Matches 12; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

Qy 3 SEAEGTFPLSTFLGIGASRLRSV 25  
Db 579 SYAGGKFPLSTYL--ADQVRSRM 598

RESULT 13  
S59116  
Programmed cell death suppressor dad-1 - *Caenorhabditis elegans*  
N;Alternative names: Gene defenser against cell death protein  
C;Species: *Caenorhabditis elegans*  
C;Date: 15-Feb-1996 #sequence revision 01-Mar-1996 #text\_change 09-Jul-2004  
C;Accession: S59116; T32142; S57533  
R;Sugimoto, A.; Hozak, R.R.; Nakashima, T.; Rotman, J.H.  
EMBO J. 14, 4414-4441, 1995  
A;Title: dad-1, an endogenous programmed cell death suppressor in *Caenorhabditis elegans*  
A;Reference number: S59116; MUID:96003623; PMID:7556086  
A;Accession: S59116  
A;Molecule type: mRNA  
A;Residues: 1-113 <SUG>  
A;Cross-references: UNIPROT:P52872; EMBL:X88080; NID:9887395; PIDN:CAA61451.1; PID:98873  
R;Greco, T.; Elliott, G.; Keppler, D.  
Submitted to the EMBL Data Library, December 1997  
A;Description: The sequence of *C. elegans* cosmid F57B10.  
A;Reference number: Z21219  
A;Accession: T32742  
A;Species: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-113 <CRE>  
A;Cross-references: EMBL:AF029713; PIDN:AB086727.1; GSDB:GN00019; CESP:F57B10.10  
A;Experimental source: strain Bristol N2; clone F57B10  
A;Introns: 71/1  
A;Map position: 1  
C;Superfamily: apoptotic cell death regulator DAD1  
C;Keywords: apoptosis

Query Match Score 46; DB 2; Length 113;  
Best Local Similarity 50.0%; Pred. No. 5.9;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 7 GTFPPLSTFLGIGASRLRS 24  
Db 49 GTFPENSFLSGFISRTWS 66

RESULT 14  
AE0336  
probable membrane protein YPO0677 [Imported] - *Yersinia pestis* (strain CO92)  
C;Species: *Yersinia pestis*  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AB0336  
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G./  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 52-527, 2001  
A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AB0326  
A;Species: preliminary  
A;Molecule type: DNA  
A;Residues: 1-126 <KUR>  
A;Cross-references: UNIPROT:Q8ZDB1; GB:AL590842; PIDN:CAC92916.1; PID:915980656; GSPPDB:G  
C;Genetics:  
A;Gene: YPO2677  
C;Superfamily: hypothetical protein MJ1523



[2] RN SEQUENCE FROM N.A.  
 RP Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;  
 RA Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; AY45648; AA; 44843 MW; E763DC581B810193 CRC64;  
 SQ SEQUENCE 402 AA; 4202 MW; 9655B3DAD962D69F CRC64;  
 Query Match 42.3%; Score 52; DB 2; Length 402;  
 Best Local Similarity 52.6%; Pred. No. 16;  
 Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 GVSEABGTPLSTFLIGIA 19  
 Db 36 GATASDGQQLLSTFLIGIA 54

RESULT 3  
 Q6SGC7 PRELIMINARY; PRT; 429 AA.  
 AC Q6SGC7; 27, Created  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TRIMBLrel. 27, Last annotation update)  
 DE Drug resistance transporter, Bcr/CfIA family.  
 DR ORFNAMES=EPAC00047A08.14;  
 OS uncultured bacterium 561.  
 OC Bacteria; environmental samples.  
 RN [1] NCBI\_TaxID=257396;  
 RP SEQUENCE FROM N.A.  
 RA DeLong E.F.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.  
 RN [2] SEQUENCE FROM N.A.  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AY418643; AACR27935.1; -  
 DR InterPro: IPR004812; EEF1UX\_Bcr\_CfIA.  
 DR TIGR007114; IPR007114; MFS.  
 DR TIGR007010; EEF1UX\_Bcr\_CfIA; 1.  
 DR PROSITE; PS050850; MFS; 1.  
 SQ SEQUENCE 429 AA; 44843 MW; E763DC581B810193 CRC64;  
 Query Match 42.3%; Score 52; DB 2; Length 429;  
 Best Local Similarity 52.6%; Pred. No. 17;  
 Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 GVSEABGTPLSTFLIGIA 19  
 Db 63 GATASDGQQLLSTFLIGIA 81

RESULT 4  
 AAR37935 PRELIMINARY; PRT; 429 AA.  
 AC AAR37935; 27, Created  
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TRIMBLrel. 27, Last annotation update)  
 DE Drug resistance transporter, Bcr/cfIA family.  
 GN EBAC00047H08.14  
 OS uncultured bacterium 561.  
 OC Bacteria; environmental samples.  
 RN [1] NCBI\_TaxID=257396;  
 RP SEQUENCE FROM N.A.  
 RA "Monterey Bay Coastal Ocean Microbial Observatory environmental clone sequencing";  
 RT Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.  
 RL SEQUENCE FROM N.A.  
 RN [2] SEQUENCE FROM N.A.  
 RP Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;

RESULT 5  
 Q9ZNM3 PRELIMINARY; PRT; 705 AA.  
 AC Q9ZNM3; 10, Created  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRIMBLrel. 10, Last annotation update)  
 DE Iuta;  
 GN Name=iuta;  
 OS Vibrio orientalis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibriaceae; Vibrio.  
 OX NCBI\_TaxID=28175;  
 RN [1] SEQUENCE FROM N.A.  
 RA Murakami K., Fue H., Takimura O., Inoue H., Yamaoka Y.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.  
 CC [1]- SUBCELLULAR LOCATION: Outer membrane (By similarity).  
 DR EMBL; AB010890; BAA74703.1; -  
 DR GO; GO:0004867; C:outer membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR InterPro: IPR000531; TonB receptor.  
 DR InterPro: IPR010917; TonB receptor C.  
 DR InterPro: IPR010105; TonB-sdp\_recept.  
 DR Pfam; PF000593; TonB dep. Rec.; 1.  
 DR TIGR01783; TIGR01783; TonB-siderophor; 1.  
 DR PROSITE; PS01156; TonB\_DEPENDENT\_REC; 1.  
 XW Membrane; Outer membrane; Receptor\_TonB box.  
 SQ SEQUENCE 705 AA; 77907 MW; 440CB4C074E44212 CRC64;  
 Query Match 41.5%; Score 51; DB 2; Length 705;  
 Best Local Similarity 60.0%; Pred. No. 39;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Qy 6 EGTFPUSLFLIGIAS 20  
 Db 5 KGSFPUSLFLIGVA 19

RESULT 6  
 Q97088 PRELIMINARY; PRT; 209 AA.  
 AC Q97088; 18, Created  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TRIMBLrel. 18, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderdLocusNames=SS03112;  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OX Sulfolobus.  
 RN [1] SEQUENCE FROM N.A.  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21332226; PubMed=11427726;  
 RX She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awayez M.J., Chan-Wieder C.C.-Y., Clausen T.G., Curtis B.A.,

RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K., Medina N., Peng X.,  
 Heikamp-de Jong I., Jeffries P., Koera C.J., Gaasterland T.,  
 Thi Ngoc H.P., Redder P., Schenk M.E., Theisraut C., Tolstrup N.,  
 Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,  
 "The complete genome of the crenarchaeon Sulfolataricus P2.",  
 Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).  
 EMBL: AB06900; AA: AY43233.1; -.  
 PIR: B90497.  
 InterPro: IPRO01279; Blactmase-like.  
 Pfam: PF00753; Lacramase\_B\_1.  
 Complete proteome; Hypothetical protein.  
 Sequence 209 AA; 23000 MW; A8863D463AB056A6 CRC64;  
 SQ OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 RN [1] \_TAXID=265669;  
 RP SEQUENCE FROM N.A.  
 RX PubMed:1515801; DOI:10.1093/nar/gth562;  
 RA Nelson K.E., Routs D.E., Mongodin E.F., Ravel J., DeBoy R.T.,  
 RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,  
 RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,  
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,  
 RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,  
 RA Forberger H.A., Tran B., Kathansky J.B., Wonderling L.D., Ulrich G.A.,  
 RA Bayles D.O., Luchansky J.B., Fraser C.M.,  
 DR "Whole genome comparisons of serotype 4b and 1/2a strains of the food-  
 born pathogen *Listeria* monocytogenes reveal new insights into the food-  
 borne pathogen *Listeria* monocytogenes reveal new insights into the food-  
 core genome components of this species.",  
 RA RMBL: AE017331; AA:AT05475.1; -.  
 DR EMBL: AE017331; AA:AT05475.1; -.  
 DR GO: GO:0016740; F:transf erase activity; IEA.  
 DR InterPro: IPR000182; GCN5Acetyltrans.  
 DR Pfam: PF00583; Acetyltransferase\_1.  
 KW Complete proteome; Transf erase.  
 SEQUENCE 153 AA; 17480 MW; A4D231EC38ED5CCF CRC64;

Query Match 38.6%; Score 47.5; DB 2; Length 153;  
 Best Local Similarity 38.1%; Pred. No. 32; Mismatches 2;  
 Matches 16; Conservative 2; Indels 19; Gaps 2;

Qy 1 GVSEAGTFPLSTFL----LGIASRLRSVA 26  
 Db 62 GVNQIKRDPSSSTFLINEKDGLLKRASSMA 92

RESULT 7  
 ID Q9YFW3 PRELIMINARY; PRT; 154 AA.  
 AC 09YFW3;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Hypothetical protein;APPE0147.  
 GN Order=EDLocusName=APPE0147;  
 OS Aeropyrum pernix  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfuroccales;  
 OC Desulfurococcaceae; Aeropyrum.  
 NCBI\_TaxID=56636;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=KL1;  
 RX SEQUENCE; PubMed:10382966;  
 RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin K., Takahashi M., Sekine M., Barai S.-I., Ankai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishizawa H.,  
 RA Takamiya M., Masuda S., Funanashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushida N., Ouchi A., Aoi K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 crenarchaeon, *Aeropyrum pernix* Kl.",  
 RL DNA Res. 6:83-101 (1999).  
 DR EMBL: AP00058; BAA19056.1; -.  
 DR PIR: H72679; H72669.  
 KW Complete proteome; Hypothetical protein.  
 SEQUENCE 154 AA; 17209 MW; 1D0C1270DA1DBB7 CRC64;

Query Match 39.0%; Score 48; DB 2; Length 154;  
 Best Local Similarity 42.3%; Pred. No. 27; Mismatches 4; Indels 11; Gaps 0;  
 Matches 11; Conservative 4; Indels 11; Gaps 0;

Qy 1 GVSEAGTFPLSTFLGIASRLRSVA 26  
 Db 53 GVVDAADGYIPGENEVGLASTALGVA 78

RESULT 8  
 ID Q71W41 PRELIMINARY; PRT; 153 AA.  
 AC Q71W41;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Acetyltransferase; GNAT family.  
 GN Order=EDLocusName=MNOF2365.270;  
 OS *Listeria* monocytogenes (serotype 4b / strain F2365).

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 RN [1] \_TAXID=265669;  
 RP SEQUENCE FROM N.A.  
 RX PubMed:1515801; DOI:10.1093/nar/gth562;  
 RA Nelson K.E., Routs D.E., Mongodin E.F., Ravel J., DeBoy R.T.,  
 RA Kolonay J.F., Rasko D.A., Angiuoli S.V., Gill S.R., Paulsen I.T.,  
 RA Peterson J.D., White O., Nelson W.C., Nierman W.C., Beanan M.J.,  
 RA Brinkac L.M., Daugherty S.C., Dodson R.J., Durkin A.S., Madupu R.,  
 RA Haft D.H., Selengut J., Van Aken S.E., Khouri H.M., Fedorova N.,  
 RA Forberger H., Tran B., Kathansky J.B., Wonderling L.D., Ulrich G.A.,  
 RA Bayles D.O., Luchansky J.B., Fraser C.M.,  
 DR "Whole genome comparisons of serotype 4b and 1/2a strains of the food-  
 borne pathogen *Listeria* monocytogenes reveal new insights into the food-  
 borne pathogen *Listeria* monocytogenes reveal new insights into the food-  
 core genome components of this species.",  
 RA RMBL: AE017331; AA:AT05475.1; -.  
 DR EMBL: AE017331; AA:AT05475.1; -.  
 DR InterPro: IPR000182; GCN5Acetyltrans.  
 DR Pfam: PF00583; Acetyltransferase\_1.  
 KW Complete proteome; Transf erase.  
 SEQUENCE 153 AA; 17480 MW; A4D231EC3BED5CCF CRC64;

Query Match 38.6%; Score 47.5; DB 2; Length 153;  
 Best Local Similarity 38.1%; Pred. No. 32; Mismatches 2;  
 Matches 16; Conservative 2; Indels 19; Gaps 2;

Qy 4 EAEGTFPLSTFL----LIGIAS-----RLRSVA 26  
 Db 34 DMEGTFHLGAFEKDVLIGIASFYPEKSTVNNPAQTRIGVA 75

RESULT	SQ	SEQUENCE	Length	DB	Score	Indels	Gaps
10	Q9M033	PRELIMINARY; PRT; 1467 AA.	1467	AA.	0;	0;	0;
	AC	09M033; PRELIMINARY;					
	DT	01-OCT-2000 (TREMBrel. 15, Created)					
	DT	01-OCT-2000 (TREMBrel. 15, Last sequence update)					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	GN	Name-T1008_110;					
	OS	Arabidopsis thaliana (Mouse-ear cress);					
	OC	Spermatophyta; Magnoliophyta; Streptophyta; Embryophyta; Tracheophyta; OC					
	OC	eudicots; rosids; OC					
	OC	Brassicaceae; Brassicales; OC					
	NCBI_TaxID=3702;	Arabidopsis; Arabidopsis.					
	RN	[1]					
	SEQUENCE FROM N.A.						
	RA	Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S., Lemke K., Mayer K.F.X.;					
	RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.					
	RN	[2]					
	SEQUENCE FROM N.A.						
	RA	EU Arabidopsis sequencing project; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.					
	RL	EMBL; AL11746; CAB81923.1; -.					
	DR	PIR; T48102; T48162.					
	DR	InterPro; IPR008938; ARM.					
	KW	Hypothetical protein.					
	SEQUENCE	1467 AA; 159477 MW; 414CA3C4DD9EA705 CRC64;					
	SQ						
	Query Match	38.6%; Score 47.5%; DB 2; Length 1467;					
	Best Local Similarity	58.3%; Pred. No. 2.9e+02; 7;					
	Matches	14; Conservative 2; Mismatches 7; Indels 1; Gaps 1;					
	Qy	2 VSEAGTTFP-LSTFLIGASRLRS 24					
	Db	1146 VLQAGAFAFPALSDFFILEILSRLVS 1169					
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport protein.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0016020; Cmembrane; IEA.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation of the substrate across the membrane (By similarity).					
	CC	-- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).					
	CC	-- SIMILARITY: Belongs to the binding-protein-dependent transport system permease family.					
	DR	EMBL; X66092; CAA46887.1; -.					
	DR	PIR; A56641; A56641.					
	DR	GO; GO:0005215; Etransporter activity; IEA.					
	DR	GO; GO:0006180; Ptransport; IEA.					
	DR	InterPro; IPR000515; BPD transp.					
	DR	Pfam; PF0528; BPD transp_1; 1.					
	DR	PROSITE; PS50328; ABC-TMI_1.					
	DR	Transmembrane Transport.					
	KW						
	RN	[1]					
	SEQUENCE FROM N.A.						
	STRAIN	=NCYC8239;					
	MEDLINE	=33113001; PubMed=1472712;					
	RA	Holck A.L., Blom H.;					
	RT	"The nucleotide sequence of a putative membrane transport gene from Clostridium perfringens."					
	DT	01-MAR-2004 (TREMBrel. 26, Last annotation update)					
	CC	DNA Seq. 3:191-194(1992).					
	CC	-1- FUNCTION: Part of a binding-protein-dependent transport system.					
	CC	Probably responsible for the translocation					

[2] RN  
SEQUENCE FROM N.A.  
RP  
RA Bueli R.;  
RL Submitted (NOV 2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AC051749; AAU31076.1;  
DR Gramene; Q8W5D1; -.  
KW Polyprotein.  
FT NON\_TBR 1  
SQ SEQUENCE 491 AA; 1 55838 MW; 3C18C5544FCEFB397 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 506;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 FPLSTFLIGIASRL 22  
Db 23 PPLTFLAGKASRL 36

---

RESULT 15  
Q6BTKL PRELIMINARY;  
PRT; 532 AA.  
ID Q6BTKL PRELIMINARY;  
AC 6BTKL;  
DT 01-OCT-2004 (TREMBLref). 28, Created  
DT 01-OCT-2004 (TREMBLref). 28, Last annotation update  
DE Similar to tr|Q8X1Y6 Debaryomyces occidentalis SCR1 protein.  
GN ORFNames=DEFA0191189;  
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomyces.  
OC Saccharomycetales; Saccharomyctaceae; Debaryomyces.  
OX NCBI\_TaxID=4959;  
RN [1]

SEQUENCE FROM N.A.  
RC STRAIN=BS767;  
RG GENOLEVURES  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marcik C., Neuveglise C., Talia E.,  
RA Goffard N., Frangeul L., Aigle M., Babour A., Bleykasten C.,  
RA Barray S., Blanchin S., Bercerich J.M., Beyne E., Boileux A., Barbe V.,  
RA Boisrame A., Boyer J., Cattolico L., Concanieri P., de Darvuar A.,  
RA Despons L., Fabre E., Faucheur C., Ferry-Dumazet H., Groppi A.,  
RA Hantayre F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszuł R., Lemaire M., Lusur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Portier S., Richard G.F., Straub M.L., Suleau A.,  
RA Swenane D., Texier F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Rukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpaelli C., Gaillard C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT Genome evolution in yeasts.";  
RN Nature 430:35-44 (2004).

[2]  
SEQUENCE FROM N.A.  
RN MEDLINE=93252821; PubMed=7334965;  
RA Mews H.W., Mayer K.P.X., Lemcke K., Schueler C.,  
RT "Directed tagging of the Arabidopsis Party ACID ELONGATION (PAE1)  
gene with the maize transposon activator.", Plant Cell 7:309-319 (1995).  
RN [2]

RP SEQUENCE FROM N.A.  
RA Bevan M., Ardiles W., Buysshaert C., Dasseville R.,  
RA De Clerck R., De Keyser A., Neyt P., Van Den Daele H.,  
RA Villaroel R., Gielen J., Van Montagu M., Jeese T., Heijnen L., Vos P.,  
RA Hoheisel J., Mews H.W., Mayer K.P.X., Lemcke K., Schueler C.,  
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RA EU Arabidopsis Sequencing project;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RN [4]

RP SEQUENCE FROM N.A.  
RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,  
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,  
RA Gielen J., Van Montagu M., Mewis H.W., Lemcke K., Mayer K.P.X.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
RN [5]

RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing Project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U29142; AAA70154.1; -  
DR EMBL; AL023094; CRA1831.1; -  
DR EMBL; AL161585; CAB0169.1; -  
DR PIR; T05272; T05272;  
DR GO; GO:0008415; :Fatty acid transferase activity; IEA.  
DR GO; GO:0016740; :Fatty acid transferase activity; IEA.  
DR GO; GO:0003058; P:biogenesis; IEA.  
DR InterPro; IPR001099; NC\_Synthase.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1  
SQ SEQUENCE 506 AA; 56263 MW; 4516D0EF8E453D18 CRC64;

Query Match 38.2%; Score 47; DB 2; Length 506;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 FPLSTFLIGIASRL 22  
Db 23 PPLTFLAGKASRL 36

---

RESULT 15  
Q6BTKL PRELIMINARY;  
PRT; 532 AA.  
ID Q6BTKL PRELIMINARY;  
AC 6BTKL;  
DT 01-OCT-2004 (TREMBLref). 28, Created  
DT 01-OCT-2004 (TREMBLref). 28, Last annotation update  
DE Similar to tr|Q8X1Y6 Debaryomyces occidentalis SCR1 protein.  
GN ORFNames=DEFA0191189;  
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomyces.  
OC Saccharomycetales; Saccharomyctaceae; Debaryomyces.  
OX NCBI\_TaxID=4959;  
RN [1]

SEQUENCE FROM N.A.  
RC STRAIN=BS767;  
RG GENOLEVURES  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marcik C., Neuveglise C., Talia E.,  
RA Goffard N., Frangeul L., Aigle M., Babour A., Bleykasten C.,  
RA Barray S., Blanchin S., Bercerich J.M., Beyne E., Boileux A., Barbe V.,  
RA Boisrame A., Boyer J., Cattolico L., Concanieri P., de Darvuar A.,  
RA Despons L., Fabre E., Faucheur C., Ferry-Dumazet H., Groppi A.,  
RA Hantayre F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszuł R., Lemaire M., Lusur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Portier S., Richard G.F., Straub M.L., Suleau A.,  
RA Swenane D., Texier F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Rukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpaelli C., Gaillard C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT Genome evolution in yeasts.";  
RN Nature 430:35-44 (2004).

[2]  
SEQUENCE FROM N.A.  
RN MEDLINE=93252821; PubMed=7334965;  
RA Mews H.W., Mayer K.P.X., Lemcke K., Schueler C.,  
RT "Directed tagging of the Arabidopsis Party ACID ELONGATION (PAE1)  
gene with the maize transposon activator.", Plant Cell 7:309-319 (1995).  
RN [2]

RP SEQUENCE FROM N.A.  
RA Bevan M., Ardiles W., Buysshaert C., Dasseville R.,  
RA De Clerck R., De Keyser A., Neyt P., Van Den Daele H.,  
RA Villaroel R., Gielen J., Van Montagu M., Jeese T., Heijnen L., Vos P.,  
RA Hoheisel J., Mews H.W., Mayer K.P.X., Lemcke K., Schueler C.,  
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RA EU Arabidopsis Sequencing project;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.  
RN [4]

RP SEQUENCE FROM N.A.  
RA Terry N., Ardiles W., Buysshaert C., Dasseville R., De Clerck R.,  
RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villaroel R.,  
RA Gielen J., Van Montagu M., Mewis H.W., Lemcke K., Mayer K.P.X.,  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
RN [5]

RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing Project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U29142; AAA70154.1; -  
DR EMBL; AL023094; CRA1831.1; -  
DR EMBL; AL161585; CAB0169.1; -  
DR PIR; T05272; T05272;  
DR GO; GO:0008415; :Fatty acid transferase activity; IEA.  
DR GO; GO:0016740; :Fatty acid transferase activity; IEA.  
DR GO; GO:0003058; P:biogenesis; IEA.  
DR InterPro; IPR001099; NC\_Synthase.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1  
SQ SEQUENCE 506 AA; 56263 MW; 4516D0EF8E453D18 CRC64;

Search completed: November 10, 2004, 13:38:51  
Job time : 34.7478 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

Copyright (c) 1993 - 2004 Compugen Ltd.  
GenCore version 5.1.6

OM protein - protein search, using sw model  
run on: November 10, 2004, 12:32:37 ; Search time 10.7081 Seconds

Title: US-10-092-750-43  
perfect score: 156  
191.991 Mill

Sequence: 1 RAPRFIXQIILDLKREIDFNVRVLFNPLS 31

Searched: Gapext 10.0 , Gapext 0.5  
Searched: 478139 secs. 66318000 residues

卷之三

卷之三

Minimum DB seq length: 0

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

```

Issue#   Patient's ID: 1/iaa/5A COMB.pep:*
1: /cn2/6/ptodatta/1/iaa/5A COMB.pep:*
2: /cn2/6/ptodatta/1/iaa/5B COMB.pep:*
3: /cn2/6/ptodatta/1/iaa/6A COMB.pep:*
4: /cn2/6/ptodatta/1/iaa/6B COMB.pep:*

```

5: /cgn2/6/pctodata1/iaa/PCTUS COMB.pep:  
6: /cgn2/6/pctodata1/iaa/backfiles1.pep:  
  
No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed. This is derived by analysis of the total score distribution.

COMMUNITIES

Result	No.	Score	Query	Length	ID	Description
			Match	DB		
1	60.5	38.8	338	4	US-09-580-929-4	Sequence 4, Appli
2	58.5	37.5	339	4	US-09-583-110-268	Sequence 3268, Ap
3	56.5	36.2	363	4	US-09-489-039A-13195	Sequence 13195, A
4	55	35.3	358	4	US-09-0248-796A-17253	Sequence 17253, A
5	53	34.0	834	2	US-08-861-644-4	Sequence 4, Appli
6	53	34.0	834	3	US-08-396-001-4	Sequence 4, Appli
7	53	34.0	834	3	US-09-323-433A-4	Sequence 4, Appli
8	53	34.0	834	4	US-09-826-754-4	Sequence 4, Appli
9	51.5	33.0	409	4	US-09-723-546-9	Sequence 9, Appli
10	50	32.1	338	4	US-09-710-279-1756	Sequence 1756, Ap
11	50	32.1	406	3	US-09-134-001C-3202	Sequence 3202, Ap
12	47.5	30.4	575	2	US-08-166-58A-5	Sequence 5, Appli
13	47	30.1	203	3	US-09-124-141-9	Sequence 9, Appli
14	47	30.1	226	3	US-09-134-001C-5494	Sequence 5494, Ap
15	47	30.1	250	3	US-09-124-141-17	Sequence 17, Appli
16	47	30.1	283	4	US-09-582-772-2	Sequence 2, Appli
17	47	30.1	376	4	US-09-540-236-2180	Sequence 2180, Ap
18	47	30.1	441	3	US-09-630-915A-34	Sequence 34, Appli
19	47	30.1	441	4	US-09-879-57-34	Sequence 34, Appli
20	47	30.1	593	3	US-09-124-141-7	Sequence 7, Appli
21	47	30.1	594	2	US-08-592-696-2	Sequence 2, Appli
22	47	30.1	594	2	US-08-592-696-4	Sequence 4, Appli
23	47	30.1	594	2	US-09-027-336-2	Sequence 2, Appli
24	47	30.1	594	2	US-09-027-336-4	Sequence 4, Appli
25	47	30.1	594	3	US-09-028-148-2	Sequence 2, Appli
26	47	30.1	594	3	US-09-028-148-4	Sequence 4, Appli
27	47	30.1	594	3	US-09-124-141-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1  
US-09-580-929-4  
; Sequence 4 , Application US/09580929  
; Patent No. 658910  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Joseph S.  
; ATTORNEY: Creuzen, Carole  
; TITLE OF INVENTION: WPPB and Net  
; FILE REFERENCE: 650-189  
; CURRENT APPLICATION NUMBER: US/09-  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 60/  
; PRIOR FILING DATE: 1999-05-28  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: E. Coli  
US -09-580-929-4

```

Query Match          64.2%;  Accn: 0.33%;  Dsgn: 33%;  Gaps 1;
Best Local Similarity 46.2%;  Pred: 0.33%;  Dsgn: 33%;  Gaps 1;
Matches 12;  Conservative 8;  Mismatches 5;  Indels 1;  Gaps 1;

Qy      6 IKQLLDL-KRELDENVRLVEFNPL 30
        ::||| | | : ||:: ||| | |:
Db      156 VEQQLTDIQLKAQQDWSTALLRENVV 181

RESULT 2
US-09-583-110-3268
; Sequence 3268, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583.110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3268
; LENGTH: 3268

```

RESULT 3  
 US-09-489-039A-13195  
 ; TYPE: PRT      ORGANISM: *Streptococcus pneumoniae*  
 US-09-583-110-3268  
 ;  
 Query Match      Score 58.5; DB 4; Length 339;  
 Best Local Similarity 38.7%; Pred. No. 0.66;  
 Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;  
 Qy 1 RAPRFKQILDL-KREIDENRVLVYFNPL 30  
 Db 152 RYKLMLESLTDIYKADSEAVNVLRLYFNPL 182

RESULT 3  
 US-09-489-039A-13195  
 ; Sequence 13195, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Bretton et. al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; CURRENT FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 13195  
 ; LENGTH: 363  
 ; TYPE: PRT  
 ; ORGANISM: *Klebsiella pneumoniae*

US-09-489-039A-13195  
 Query Match      Score 56.5; DB 4; Length 363;  
 Best Local Similarity 42.3%; Pred. No. 1.4;  
 Matches 11; Conservative 9; Mismatches 5; Indels 1; Gaps 1;  
 Qy 6 IKQILDL-KREIDENRVLVYFNPL 30  
 Db 181 VEQILDLQKQPEWSTALERYFNPV 206

RESULT 4  
 US-09-248-796A-17253  
 ; Sequence 17253, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 101196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; CURRENT FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; PRIOR FILING DATE: 1998-08-13  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; LENGTH: 358  
 ; TYPE: PRT  
 ; ORGANISM: *Candida albicans*

US-09-248-796A-17253  
 Query Match      Score 55; DB 4; Length 358;  
 Best Local Similarity 45.8%; Pred. No. 2.4;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

RESULT 5  
 US-08-861-464-4  
 ; Sequence 4, Application US/08861464  
 ; Patent No. 5874210  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Guarante, Leonard P.  
 ; APPLICANT: Austriaco Jr., Nicanor  
 ; APPLICANT: Kennedy, Brian  
 ; TITLE OF INVENTION: Genes Determining Cellular Senescence  
 ; TITLE OF INVENTION: in Yeast  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 ; STREET: Two Militia Drive  
 ; CITY: Lexington  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02113  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/861,464  
 ; FILING DATE: 22-MAY-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/396,001  
 ; FILING DATE: 28-FEB-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/09351  
 ; FILING DATE: 15-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/107,408  
 ; FILING DATE: 16-AUG-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Granahan, Patricia  
 ; REGISTRATION NUMBER: 32,227  
 ; REFERENCE/DOCKET NUMBER: MTT-6408A22  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 781-861-6240  
 ; TELEX/FAX: 781-861-9540  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 834 amino acids  
 ; TYPE: amino acid  
 ; TOPology: linear  
 ; MOLECULE TYPE: protein  
 US-08-861-644-4

Query Match      Score 53; DB 2; Length 834;  
 Best Local Similarity 55.6%; Pred. No. 1.3;  
 Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

Qy 5 FIKQILDLKREIDENRVLVYFNPLS 31  
 Db 392 YIQFLLDIK-ELDFYL-LAFLFNRS 415

RESULT 6  
 US-08-396-001-4  
 ; Sequence 4, Application US/08396001  
 ; Patent No. 5919618  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Guarante, Leonard P.  
 ; APPLICANT: Austriaco Jr., Nicanor  
 ; APPLICANT: Claus, James  
 ; APPLICANT: Coles, Francesca  
 ; APPLICANT: Kennedy, Brian  
 ; TITLE OF INVENTION: Genes Determining Cellular Senescence in  
 ; YEAST  
 ; NUMBER OF SEQUENCES: 16

Qy 1 RAPRFKQILDLKREIDENRVLV 24  
 Db 179 RANTFISQQLDQTFNSDFNINVV 202

CORRESPONDENCE ADDRESS:  
ADDRESSSE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS-MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.3.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/396,001  
FILING DATE: 28-FEB-1995  
CLASIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REFERENCE/DOCKET NUMBER: MIT-6408A2  
TELEPHONE: 617-861-2240  
TELEX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 4 :  
SEQUENCE CHARACTERISTICS:  
LENGTH: 834 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-396-001-4

Query Match 34.0%; Score 53; DB 2; Length 834;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

Qy 5 FIKQILDLKREIDENVRLEYFNPLS 31  
Db 392 YIIQFLLDIK-ELDFYL-LAEELFRLS 416

RESULT 7  
US-09-323-433A-4

Sequence 4, Application US/09323433A  
PATENT NO. 6218512

GENERAL INFORMATION:

APPLICANT: Guarante, Leonard P.  
APPLICANT: Austraciaco Jr., Nicanor  
APPLICANT: Claus, James J.  
APPLICANT: Cole, Francesca  
APPLICANT: Kennedy, Brian

TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN

FILE REFERENCE: 0050-1491-003  
CURRENT APPLICATION NUMBER: US/09/433A  
CURRENT FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: US 08/395,001  
PRIOR FILING DATE: 1995-02-28

PRIOR APPLICATION NUMBER: PCT/US94/09331  
PRIOR FILING DATE: 1994-08-15  
PRIOR APPLICATION NUMBER: US 08/107,408  
PRIOR FILING DATE: 1993-08-16

PRIOR APPLICATION NUMBER: US 09/323,433  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSEQ For Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 834  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-323-433A-4

Query Match 34.0%; Score 53; DB 2; Length 834;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

RESULT 8  
US-09-826-752-4

Sequence 4, Application US/09826752  
PATENT NO. 6787100  
GENERAL INFORMATION:  
APPLICANT: Guarante, Leonard P.  
APPLICANT: Austraciaco Jr., Nicanor  
APPLICANT: Claus, James J.  
APPLICANT: Cole, Francesca  
APPLICANT: Kennedy, Brian

TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN

FILE REFERENCE: 0050-1491-005  
CURRENT APPLICATION NUMBER: US/09/826,752  
CURRENT FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: US 08/396,001  
PRIOR FILING DATE: 1995-02-28  
PRIOR APPLICATION NUMBER: PCT/US94/09351  
PRIOR FILING DATE: 1994-08-15  
PRIOR APPLICATION NUMBER: US 08/107,408  
PRIOR FILING DATE: 1993-08-16  
PRIOR APPLICATION NUMBER: US 09/323,433  
PRIOR FILING DATE: 1999-06-01  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: FastSEQ For Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 834  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-826-752-4

Query Match 34.0%; Score 53; DB 4; Length 834;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 15; Conservative 4; Mismatches 6; Indels 2; Gaps 1;  
Qy 5 FIKQILDLKREIDENVRLEYFNPLS 31  
Db 392 YIIQFLLDIK-ELDFYL-LAEELFRLS 416

RESULT 9  
US-09-723-546-9

Sequence 9, Application US/09723546  
PATENT NO. 6737259  
GENERAL INFORMATION:  
APPLICANT: Clark, Mike A.  
APPLICANT: Modified Arginine Deiminase

TITLE OF INVENTION: Modified Arginine Deiminase

FILE REFERENCE: PHOEO59  
CURRENT APPLICATION NUMBER: US/09/723,546  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: 09/023,809  
PRIOR FILING DATE: 1998-02-13  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 409  
TYPE: PRT  
ORGANISM: Borrelia afzelii  
US-09-723-546-9

Query Match 33.0%; Score 51.5; DB 4; Length 409;  
Best Local Similarity 34.5%; Pred. No. 9.6;  
Matches 10; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

Qy 4 RFIKQILDLKREIDENVRLEYFNPLS 31  
Db 89 KFISQFILEABIKTDEFLNLKDIFSSLT 117

Query Match 34.0%; Score 53; DB 3; Length 834;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 15; Conservative 4; Mismatches 6; Indels 2;

RESULT 10  
US-09-710-279-1756 Application US/09710279  
Sequence 1756 Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PU3400US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SEQ ID NO: 1756  
SEQUENCE: PatentIn Ver. 2.1  
LENGTH: 3388  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-1756

Query Match 32.1%; Score 50; DB 4; Length 338;  
Best Local Similarity 31.2%; Pred. No. 13;  
Matches 10; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

Qy 5 FIKQILDLRE-----IDFNVLVLFVN 28  
Db 39 FVRDLYQDVQRDDFTMNAIMDNLYRLDVFN 70

RESULT 11  
US-09-134-001C-3202  
Sequence 3202 Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIORITY APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO: 3202  
LENGTH: 406  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3202

Query Match 32.1%; Score 50; DB 3; Length 406;  
Best Local Similarity 31.2%; Pred. No. 16;  
Matches 10; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

Qy 5 FIKQILDLRE-----IDFNVLVLFVN 28  
Db 107 FVRDLYQDVQRDDFTMNAIMDNLYRLDVFN 138

RESULT 12  
US-08-766-858A-5  
Sequence 5 Application US/08766858A  
Patent No. 595782  
GENERAL INFORMATION:  
APPLICANT: Cianciotto, Nicholas P.  
APPLICANT: Hickey, Erin K.  
APPLICANT: O'Connell, William A.  
TITLE OF INVENTION: METHODS AND MATERIALS FOR DETECTING  
OTHER INFORMATION: Description of unknown Organism: N-terminal  
US-09-124-141-9

Query Match      Score 47; DB 3; Length 203;  
 Best Local Similarity 35.5%; Pred. No. 20;  
 Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;  
 Qy      3 PRFIQILDLKREIDDNVRL--VEYFNPLS 31  
 Db      162 PESLQILVDCRDTLYKGVRTGHPRFFNQLS 192

Search completed. November 10, 2004, 13:44:12  
 Job time : 11.7081 secs

RESULT 14  
 US-09-134-001C-5494  
 Sequence 5494, Application US/09134001C  
 GENERAL INFORMATION:  
 APPLICANT: Lynn Doucette-Stamm et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 CURRENT APPLICATION NUMBER: US/09/134,001C  
 CURRENT FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: US 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: US 60/055,779  
 PRIOR FILING DATE: 1997-08-14  
 NUMBER OF SEQ ID NOS: 5674  
 SEQ ID NO 5494  
 LENGTH: 226  
 TYPE: PRT  
 ORGANISM: *Staphylococcus epidermidis*  
 US-09-134-001C-5494

Query Match      Score 47; DB 3; Length 226;  
 Best Local Similarity 41.7%; Pred. No. 23;  
 Matches 15; Conservative 5; Mismatches 4; Indels 12; Gaps 2;  
 Qy      5 FIKQIL--DKEI--DENYVLEYEN 28  
 Db      1 FVMQTLLVEDDDNTLFQELKKELEQMDENYVGVEDFS 36

RESULT 15  
 US-09-124-141-17  
 Sequence 17, Application US/09124141  
 GENERAL INFORMATION:  
 APPLICANT: Harrison, Leonard  
 APPLICANT: Honeyman, Margot  
 APPLICANT: Cram, David  
 APPLICANT: De Aipurua, Henry  
 TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC  
 FILE REFERENCE: Phillips, Ormonde & Fitzpatrick  
 CURRENT APPLICATION NUMBER: US/09/124,141  
 CURRENT FILING DATE: 1998-07-29  
 EARLIER FILING DATE: 1994-09-20  
 EARLIER APPLICATION NUMBER: 07/839,805  
 EARLIER FILING DATE: 1992-02-21  
 NUMBER OF SEQ ID NOS: 34  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 17  
 LENGTH: 250  
 TYPE: PRT  
 ORGANISM: Unknown Organism  
 FEATURE: OTHER INFORMATION: Description of Unknown Organism:N-terminal  
 OTHER INFORMATION: Fragment of Human Brain GAD (HBGAD17)

Query Match      Score 47; DB 3; Length 250;  
 Best Local Similarity 35.5%; Pred. No. 26;  
 Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;  
 US-09-124-141-17

This Page Blank (uspto)

Copyright Gencore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW mode!

Run on: November 10, 2004, 16:36:12 ; Search time 33:3022 Seconds

(without alignments)  
328.807 Million cell updates/sec

Title: US-10-092-750-43

Perfect score: 156

Sequence: 1 RAPRFIQIQLLDLKREIDFNVLVYFNPDS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 15666620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 15666620

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : Published Applications AA,\*

1: /cgn2\_6\_ptodata/1/pubbaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6\_ptodata/1/pubbaa/PC7\_NEW\_PUB.pep:\*

3: /cgn2\_6\_ptodata/1/pubbaa/US07\_PUB.pep:\*

4: /cgn2\_6\_ptodata/1/pubbaa/US07\_PUBCOMB.pep:\*

5: /cgn2\_6\_ptodata/1/pubbaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6\_ptodata/1/pubbaa/PC07\_PUBCOMB.pep:\*

7: /cgn2\_6\_ptodata/1/pubbaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6\_ptodata/1/pubbaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6\_ptodata/1/pubbaa/US09A\_PUBCOMB.pep:\*

10: /cgn2\_6\_ptodata/1/pubbaa/US09B\_PUBCOMB.pep:\*

11: /cgn2\_6\_ptodata/1/pubbaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6\_ptodata/1/pubbaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6\_ptodata/1/pubbaa/US10A\_PUBCOMB.pep:\*

14: /cgn2\_6\_ptodata/1/pubbaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6\_ptodata/1/pubbaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6\_ptodata/1/pubbaa/US10D\_PUBCOMB.pep:\*

17: /cgn2\_6\_ptodata/1/pubbaa/US10\_NEW\_PUB.pep:\*

18: /cgn2\_6\_ptodata/1/pubbaa/US11\_NEW\_PUB.pep:\*

19: /cgn2\_6\_ptodata/1/pubbaa/US60\_NEW\_PUB.pep:\*

20: /cgn2\_6\_ptodata/1/pubbaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
- 1	156	100.0	31	14 US-10-092-750-43	Sequence 43, App1
2	86	55.1	167	9 US-09-864-761-34496	Sequence 34496, A
3	78.5	50.3	1031	14 US-10-082-878A-257	Sequence 257, App
4	78.5	50.3	1275	14 US-10-025-201-3	Sequence 3, App1
5	74	47.4	363	14 US-10-012-600B-221	Sequence 221, App
6	73.5	47.1	1274	16 US-10-415-615-3	Sequence 3, App1
7	72.5	44.5	1272	15 US-10-114-70-10	Sequence 2971, App
8	65.5	42.0	148	14 US-10-104-047-2971	Sequence 1442, App
9	61.5	39.4	78	9 US-09-867-550-1462	Sequence 8, App1
10	60	38.8	338	9 US-09-318-271-8	Sequence 214640,
11	58	37.2	261	15 US-10-424-599-214640	Sequence 44206, A
12	57.5	36.9	216	15 US-10-425-114-44206	Sequence 8963, A
13	57.5	36.9	351	17 US-10-739-930-8963	Sequence 8963, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
- 1	156	100.0	31	14 US-10-092-750-43	Sequence 43, App1
2	86	55.1	167	9 US-09-864-761-34496	Sequence 34496, A
3	78.5	50.3	1031	14 US-10-082-878A-257	Sequence 257, App
4	78.5	50.3	1275	14 US-10-025-201-3	Sequence 3, App1
5	74	47.4	363	14 US-10-012-600B-221	Sequence 221, App
6	73.5	47.1	1274	16 US-10-415-615-3	Sequence 3, App1
7	72.5	44.5	1272	15 US-10-114-70-10	Sequence 2971, App
8	65.5	42.0	148	14 US-10-104-047-2971	Sequence 1442, App
9	61.5	39.4	78	9 US-09-867-550-1462	Sequence 8, App1
10	60	38.8	338	9 US-09-318-271-8	Sequence 214640,
11	58	37.2	261	15 US-10-424-599-214640	Sequence 44206, A
12	57.5	36.9	351	17 US-10-739-930-8963	Sequence 8963, A

RESULT 1  
US-10-092-750-43  
; Sequence 43, Application US/10092750  
; Publication No. US003003215TA1  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; ATTORNEY: Wright, Martin C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; CURRENT APPLICATION NUMBER: 50036/05002  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: US/10/092,750  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-092-750-43

RESULT 2  
US-09-564-761-34496  
; Sequence 34496, Application US/09864761  
; Patent No. US2002048763A1  
; GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wenheng  
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 FILE REFERENCE: Aeomica-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 SOFTWARE: Annamax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 34496  
 LENGTH: 167  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC006459.2  
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.9  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6  
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 17  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.3  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 11  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5  
 OTHER INFORMATION: EXPRESSED IN SWISSPROT HIT: AWP82328.1, EVALU 4.00e-59  
 OTHER INFORMATION: SWISSPROT HIT: P08548, EVALU 1.00e-51  
 US-09-864-761-34496

Query Match 55.1%; Score 86; DB 9; Length 167;  
 Best Local Similarity 59.3%; Pred. No. 0.00016; Mismatches 6; Indels 0; Gaps 0;

APRFIKQILDLRKEIDFNRVLYFN PLS 2

Db 24 APRFIKQILDLRKEIDFNRVLYFN PLS 0

RESULT 3  
 US-10-082-828A-257  
 Sequence 257, Application US/10082828A  
 Publication No. US20030175715A1  
 GENERAL INFORMATION:  
 APPLICANT: Sun, Yongming  
 APPLICANT: Recipon, Herve  
 APPLICANT: Salceda, Susana  
 APPLICANT: Liu, Chenzhua  
 APPLICANT: Turner, Leah  
 TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pro.  
 FILE REFERENCE: DEX-0247  
 CURRENT APPLICATION NUMBER: US/10/082,828A  
 CURRENT FILING DATE: 2002-07-09  
 PRIOR APPLICATION NUMBER: US 60/243,805  
 PRIOR FILING DATE: 2000-10-27  
 NUMBER OF SEQ ID NOS: 266  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 257  
 LENGTH: 1031  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-082-828A-257

Query Match 50.3%; Score 78.5; DB 14; Length 1031;  
 Best Local Similarity 54.8%; Pred. No. 0.014; Mismatches 7; Indels 1; Gaps 1;  
 Matches 17; Conservative 7; Number of SEQ ID NOS: 31

Oy 2 APRFIKQILDLRKEIDFNRVLYFN PLS 31  
 Db 23 APRFIKQVLSDLQRDLSHTLIMGDENTPLS 53

RESULT 4  
 US-10-025-201-3  
 Sequence 3, Application US/10025201  
 Publication No. US2003003468A1  
 GENERAL INFORMATION:  
 APPLICANT: Crow, Mary K.  
 TITLE OF INVENTION: MARKERS FOR DISEASE SUSCEPTIBILITY AND TARGETS FOR THERAPY  
 FILE REFERENCE: 5983/2H567  
 CURRENT APPLICATION NUMBER: US/10/025,201  
 CURRENT FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: 60/256,673  
 PRIOR FILING DATE: 2000-12-19  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 3  
 LENGTH: 1275  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 DATABASE INFORMATION:  
 DATABASE ACCESSION NUMBER: GenBank Accession No. US2003003468A1 U09116  
 DATABASE ENTRY DATE: 1995-02-02  
 RELEVANT RESIDUES: (1)..(1275)  
 US-10-025-201-3

Query Match 50.3%; Score 78.5; DB 14; Length 1275;  
 Best Local Similarity 54.8%; Pred. No. 0.018; Mismatches 7; Indels 1; Gaps 1;  
 Matches 17; Conservative 7; Number of SEQ ID NOS: 31

Oy 2 APRFIKQILDLRKEIDFNRVLYFN PLS 31  
 Db 121 APRFIKQVLSDLQRDLSHTLIMGDENTPLS 151

RESULT 5  
 US-10-012-6003-221  
 Sequence 22, Application US/10012600B  
 Publication No. US2003006437A1

Qy 2 APRFIKQILDLRKEIDFNRVLYFN PLS 28  
 Db 121 APRFIKQVLSDLQRDLSHTLIMGDENTPLS 151

GENERAL INFORMATION:  
 APPLICANT: Sun, Yongming  
 APPLICANT: Recipon, Herve  
 APPLICANT: Chen, Sei-Yu  
 TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
 FILE REFERENCE: DEX-0265  
 CURRENT APPLICATION NUMBER: US/10/012,600B  
 CURRENT FILING DATE: 2001-11-16  
 PRIOR FILING DATE: 2000-11-06  
 NUMBER OF SEQ ID NOS: 240  
 SEQ ID NO: 221  
 LENGTH: 363  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-012-600B-221

Query Match Score 74; DB 14; Length 363;  
 Best Local Similarity 51.7%; Pred. No. 0.021; Indels 0; Gaps 0;  
 Matches 15; Conservative 7; Mismatches 7;

Qy	2 APRFIKQILDKRIBDNFVRLVEYFNPL 30
Db	120 APRFIKQVLRLQRDLNSHIVLVDGFTNL 148

RESULT 6  
 US-10-415-615-3  
 Sequence 3 Application US/10415615  
 Publication No. US/0040101943A1  
 GENERAL INFORMATION:  
 APPLICANT: INCYTE CORPORATION  
 APPLICANT: BAUGHN, Mariah R.  
 APPLICANT: YUE, Henry  
 APPLICANT: LU, Yan  
 APPLICANT: DING, Li  
 APPLICANT: TANG, Y. Tom  
 APPLICANT: GANDHI, Ameena R.  
 APPLICANT: HAFAJIA, April J.A.  
 APPLICANT: LAL, Preeti G.  
 TITLE OF INVENTION: NUCLEIC ACID MODIFICATION ENZYMES  
 FILE REFERENCE: PI-0260 USN  
 CURRENT APPLICATION NUMBER: US/10/415,615  
 CURRENT FILING DATE: 2003-04-29  
 PRIOR APPLICATION NUMBER: PCT/US01/46301  
 PRIOR FILING DATE: 2001-11-01  
 PRIOR APPLICATION NUMBER: US 60/245,458  
 PRIOR FILING DATE: 2000-11-02  
 PRIOR APPLICATION NUMBER: US 60/255,107  
 PRIOR FILING DATE: 2000-12-12  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: PERL Program  
 SEQ ID NO 3  
 LENGTH: 1274  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No: 1258887CD1

Query Match Score 73.5; DB 16; Length 1274;  
 Best Local Similarity 53.3%; Pred. No. 0.095; Indels 1; Gaps 1;  
 Matches 16; Conservative 7; Mismatches 6;

Qy	3 PRFIKQILDKRIBDNFVRLVEYFNPL 31
Db	122 PRFIKQVLRLQRDLNSHIVLVDGFTNL 151

RESULT 7  
 US-10-114-270-30

; Sequence 30, Application US/10114270  
 ; Publication No. US/0040030110A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Guo, Xiaochia  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Malynkar, Uriel M.  
 ; APPLICANT: Spytek, Kimberly A.  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Liu, Ziaohong  
 ; APPLICANT: Gusev, Vladimir Y.  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Verner, Corine  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Cormann, Linda  
 ; APPLICANT: Shenvoy, Suresh G.  
 ; APPLICANT: Pena, Carol E.A.  
 ; APPLICANT: Smithson, Glenda  
 ; APPLICANT: Burgess, Catherine E.  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Padigaru, Muraiidhara  
 ; APPLICANT: Shimkets, Richard A.  
 ; APPLICANT: Gangolli, Esha A.  
 ; APPLICANT: Taupier Jr., Raymond J.  
 ; APPLICANT: Casman, Stacie J.  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Liete, Mario W.  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Edinger, Shlomit R.  
 ; APPLICANT: Stone, David J.  
 ; APPLICANT: Macdougall, John R.  
 ; APPLICANT: Rothenberg, Mark E.  
 ; TITLE OF INVENTION: No. US/2004030110A1 Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-322C  
 ; CURRENT APPLICATION NUMBER: US/10/114,270  
 ; CURRENT FILING DATE: 2002-11-27  
 ; PRIOR APPLICATION NUMBER: 60/281,086  
 ; PRIOR FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: 60/281,136  
 ; PRIOR FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: 60/281,863  
 ; PRIOR FILING DATE: 2001-04-05  
 ; PRIOR APPLICATION NUMBER: 60/281,906  
 ; PRIOR FILING DATE: 2001-04-05  
 ; PRIOR APPLICATION NUMBER: 60/282,020  
 ; PRIOR FILING DATE: 2001-04-05  
 ; PRIOR APPLICATION NUMBER: 60/282,930  
 ; PRIOR FILING DATE: 2001-04-10  
 ; PRIOR APPLICATION NUMBER: 60/282,934  
 ; PRIOR FILING DATE: 2001-04-10  
 ; PRIOR APPLICATION NUMBER: 60/283,512  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/283,710  
 ; PRIOR FILING DATE: 2001-04-13  
 ; PRIOR APPLICATION NUMBER: 60/284,234  
 ; PRIOR FILING DATE: 2001-04-17  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 470  
 ; SEQ ID NO 30  
 ; LENGTH: 1272  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-114-270-30

Query Match Score 72.5; DB 15; Length 1272;  
 Best Local Similarity 51.6%; Pred. No. 0.13; Indels 1; Gaps 1;  
 Matches 16; Conservative 7; Mismatches 7;  
 Qy 2 APRFIKQILDKRIBDNFVRLVEYFNPLS 31  
 Db 117 APRFIKQVLRLQRDLNSHIVLVDGFTNL 147

RESULT 8  
 US-10-104-047-2971  
 ; Sequence 2971, Application US/10104047  
 ; Publication No. US2003033639A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US2003033639A1 full length cDNA  
 ; CURRENT APPLICATION NUMBER: US/104,047  
 ; CURRENT FILING DATE: 2002-03-25  
 ; PRIOR APPLICATION NUMBER:  
 ; PRIORITY FILING DATE:  
 ; NUMBER OF SEQ ID NOS: 4096  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 2971  
 ; LENGTH: 148  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-104-047-2971

Query Match 42.0%; Score 65.5; DB 14; Length 148;  
 Best Local Similarity 51.6%; Pred. No. 0.14%;  
 Matches 16; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 2 AFRFIQIILDKREIDENYVRLVEYFN-PLS 31  
 Db 103 AFRLIKQVLDKDRQRDLSHTITVGDFNMPUS 133

RESULT 9  
 US-09-867-550-1462  
 ; Sequence 1462, Application US/09867550  
 ; Patent No. US20082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad.  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20082206A1 Polynucleotides from Atherogenic Cells and  
 ; FILE REFERENCE: 21402\_013 (Curta 313)  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USSN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 1462  
 ; LENGTH: 78  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE: NAME/KEY: VARIANT  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Wherein Xaa may be any one of Ala or Arg or Asn or Asp or Cys or  
 ; OTHER INFORMATION: His or Ile or Leu or Phe or Pro or Ser or Thr or Val or  
 ; SEQ ID NO: 867-550-1462

Query Match 39.4%; Score 61.5; DB 9; Length 78;  
 Best Local Similarity 57.1%; Pred. No. 0.27%;  
 Matches 16; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

Query 1 RAPRFIQIILDKREIDENYVRLVEYFN 28  
 Db 19 RAPRIKQI-LDLRKVIDSKTKRADFN 45

Query Match 39.4%; Score 61.5; DB 9; Length 78;  
 Best Local Similarity 57.1%; Pred. No. 0.27%;  
 Matches 16; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

Query 1 RAPRFIQIILDKREIDENYVRLVEYFN 28  
 Db 22 KAKNVAEIIIDLSSNSDMAVMLRYFN 249

RESULT 12  
 US-10-425-114-44206  
 ; Sequence 44206, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Screen, Steven E.

RESULT 10  
 US-09-318-271-8  
 ; Sequence 8, Application US/09318271A  
 ; Patent No. US20020012979A1

APPLICANT: Tabaska, Jack E  
 NUMBER OF SEQ ID NOS: 732  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53313)B  
 CURRENT APPLICATION NUMBER: US10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 733128  
 SEQ ID NO 44206  
 LENGTH: 216  
 TYPE: PRT  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: 700889134\_FLI.pep  
 US-10-425-114-44206

Query Match Score 36.9%; Score 57.5%; DB 15; Length 216;

Best Local Similarity 35.5%; Pred. No. 3.1;  
 Matches 11; Conservative 10; Mismatches 9; Indels 1; Gaps 1;  
 SEQ ID NO 80 RTKLFVEEARDIQAETWRILLRYFPV 110

RESULT 13  
 US-10-739-930-8963  
 Sequence 8963, Application US/10739930  
 Publication No. US20030216190AA1  
 GENERAL INFORMATION:  
 APPLICANT: Kovacic, David K.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
 FILE REFERENCE: 38-21(53317)B  
 CURRENT APPLICATION NUMBER: US10/739,930  
 CURRENT FILING DATE: 2003-12-18  
 NUMBER OF SEQ ID NOS: 11088  
 SEQ ID NO 8963  
 LENGTH: 351  
 TYPE: PRT  
 ORGANISM: Glycine max  
 OTHER INFORMATION: Clone ID: GLYMA-23APR03-C2735\_7.P  
 US-10-739-930-8963

Query Match Score 36.9%; Score 57.5%; DB 17; Length 351;

Best Local Similarity 35.5%; Pred. No. 5.2;  
 Matches 11; Conservative 10; Mismatches 9; Indels 1; Gaps 1;  
 SEQ ID NO 160 RTKLFVEEARDIQAETWRILLRYFPV 190

RESULT 14  
 US-10-451-467A-588  
 Sequence 588, Application US/10451467A  
 Publication No. US20040161840A1  
 GENERAL INFORMATION:  
 APPLICANT: CONTRERAS, ROLAND HENRI  
 APPLICANT: EBERHARDT, INES  
 APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS  
 APPLICANT: REEKENS, RIEKA JOSEPHINA  
 TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN YEAST AND FUNGI  
 FILE REFERENCE: JAB-1667

CURRENT APPLICATION NUMBER: US10/451,467A  
 CURRENT FILING DATE: 2003-06-19  
 PRIOR APPLICATION NUMBER: EP 00870318.3  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: EP 01870002.1  
 PRIOR FILING DATE: 2001-01-04  
 PRIOR APPLICATION NUMBER: EP 01870003.9

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
 SEQ ID NO 158 RANTFSQLFDQTKPSDFNINVV 181

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Query Match Score 35.3%; Score 55%; DB 16;  
 Best Local Similarity 45.8%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

This Page Blank (uspto)

OM Protein - Protein search, using sw model  
Run on: November 10, 2004, 12:29:32 ; Search time 6.90074 Seconds  
Perfect score: 156  
Sequence: 1 RAPRFKQILIDLKREIDFNVRLLVYFNFPLS 31  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:

283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:

283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000.Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 SummariesDatabase : PIR 79.1\*  
1: pir1;\*  
2: pir2;\*  
3: pir3;\*  
4: pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARES

Result No.	Score	Query Match	Length	DB ID	Description
1	78.5	50.3	1275	2 B28096	line-1 protein ORF2 - human
2	78.5	50.3	1275	2 I38588	C:Species: Homo sapiens (man)
3	78.5	50.3	1275	2 S65824	C:Accession: B28096
4	77.5	49.7	1259	4 GNHULL	A:Title: Unit-length line-1 transcripts in human teratocarcinoma cells.
5	77.5	49.7	1280	2 B34087	A:Accession: A28096
6	60.5	38.8	338	1 XUBCUG	A:Cross-references: UNIPROT:Q15604; UNIPROT:Q9UN80; UNIPROT:Q12881; UNIPROT:O00363; UNIPROT:Q9Y5K0; UNIPROT:O00366; UNIPROT:Q8TE30; UNIPROT:O00375
7	60.5	38.8	338	2 C90727	C:Superfamily: pol polyprotein
8	60.5	38.8	338	2 D85578	Query Match 50.3%; Score 78.5%; DB 2; Length 1275;
9	58.5	37.5	339	2 B95187	Best Local Similarity 54.8%; Pred. No. 0.012; Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;
10	58.5	37.5	339	2 C9054	Qy 2 APRFIQQLLDRKDENDVPLVEYN-PLS 31
11	57.5	36.9	354	2 T10496	Db 121 APRFIQQLLDRKDENDVPLVEYN-PLS 151
12	57	36.5	285	2 S49879	RESULT 2
13	57	36.5	285	2 T12157	I38588 reverse transcriptase homolog - human retrotransposon L1
14	57	36.5	285	2 S49898	N1:Alternate names: ORF2 protein
15	56.5	36.2	337	1 A37760	C:Species: Homo sapiens (man)
16	56.5	36.2	338	1 S51328	C:Accession: 02-Jul-1996 #sequence_revision 02-Jul-1996
17	56.5	36.2	338	2 AF0528	R.Holmes, S.E.; Domrowski, B.A.; Krebs, C.M.; Boehm, C.D.; Kazazian, H.H.
18	56	36.2	347	2 T19989	A:Title: A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q p11
19	56	35.9	285	2 T12156	A:Reference number: I38587; PMID:95004577
20	55.5	35.6	500	2 S16788	A:Accession: I38588
21	55	35.3	796	2 T43782	A:Molecule type: DNA
22	54	34.6	431	2 D84779	A:Cross-references: UNIPROT:Q12881; PIDN:U09116; NID:G483914; PIDN:AAB60345.1; PID:G48393]
23	53.5	34.3	336	1 S70744	C:Superfamily: pol polyprotein
24	53	34.0	732	2 T32023	Query Match 50.3%; Score 78.5%; DB 2; Length 1275;
25	53	34.0	859	2 S64195	Best Local Similarity 54.8%; Pred. No. 0.012; Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;
26	52.5	33.7	350	2 T06526	Qy 2 APRFIQQLLDRKDENDVPLVEYN-PLS 31
27	52.5	33.7	351	2 B86261	Db 121 APRFIQQLLDRKDENDVPLVEYN-PLS 151
28	52.5	33.7	351	2 S62783	
29	52	33.3	205	2 T24749	

RESULT 3		Best Local Similarity 54.8%; Pred. No. 0.017; N-matches 7; Indels 6; Gaps 1;	
S65B24	reverse transcriptase homolog - human transposon L1.1	Qy 2 APRFIQIQLDKREIDFNRLVVEFN-PUS 31	Db 126 APRFIQIQLDKREIDFNRLVVEFN-PUS 31
C;Species: Homo sapiens (man)	C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004		
C;Accession: S65B24			
R;Dombroski, B.A.	submitted to the EMBL Data Library, January 1992	RESULT 6	
A;Description: Isolation of an active human transposable element.	XUCUG	XUCUG UDPgalactose 4-epimerase (EC 5.1.1.3.2) - Escherichia coli (strain K-12)	
A;Reference number: S65B23		N;Alternative names: UDPgalactose 4-epimerase	
A;Accession: S65B24		C;Species: Escherichia coli	
A;Status: preliminary		C;Date: 31-Dec-1990 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004	
A;Molecule type: DNA		C;Accession: S02089; S00759; S14391; G6811; B25764	
A;Residues: 1-1275 <DOM>		R;Lemaire, H.G.	
A;Cross-references: UNIPROT:Q15604; EMBL:M80340; PID:941522; PMID:941522; MUID:87040735; PMID:3022232		Submitted to the EMBL Data Library, April 1988	
C;Superfamily: pol polyprotein		A;Reference number: S02089	
		A;Accession: S02089	
		A;Molecule type: DNA	
Query Match Score 50.3%; DB 2; Length 1275;		A;Residues: 1-338 <LEM1>	
Best Local Similarity 54.8%; Pred. No. 0.012;		A;Cross-references: UNIPROT:P09147; EMBL:X06226; PIDN:941522; PMID:941522; MUID:941522; PMID:2134186	
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;		R;Lemaire, H.G.; Mueller-Hill, B.	
Qy 2 APRFIQIQLDKREIDFNRLVVEFN-PUS 31		Nucleic Acids Res. 14, 7705-7711, 1986	
Db 121 APRFIQIQLDKREIDFNRLVVEFN-PUS 151		A;Title: Nucleotide sequences of the galE gene and the galT gene of E. coli.	
		A;Reference number: S00722; MUID:87040735; PMID:3022232	
		A;Accession: S00759	
		A;Molecule type: DNA	
		A;Residues: 1-294; APRFIQIQLDKREIDFNRLVVEFN-PUS 310-338 <LEM2>	
		A;Cross-references: EMBL:X06226	
		A;Note: this sequence has been revised in reference S02089	
		R;Bernardi, F.; Bernardi, A.	
		DNA Seq. 1, 141-150, 1990	
		A;Title: Completed sequence of PKG1800, a vector for determination of transcription termi	
		A;Accession: S14391	
		A;Status: transcription not shown	
		A;Molecule type: DNA	
		A;Residues: 1-139, 'LPLIPGQ' <BER>	
		A;Cross-references: EMBL:X5149; PIDN:942412; PMID:942413; MUID:9190543; PMID:2134186	
		R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Burland, V.; Riley, M.; Co	
		A;Reference number: S14391	
		A;Status: transcription not shown	
		A;Molecule type: DNA	
		A;Residues: 1-139, 'LPLIPGQ' <BER>	
		A;Cross-references: EMBL:X5149; PIDN:942412; PMID:942413; MUID:9190543; PMID:2134186	
		R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Burland, V.; Riley, M.; Co	
		A;Reference number: S14391	
		A;Status: transcription not shown	
		A;Molecule type: DNA	
		A;Residues: 1-138 <BLAT>	
		A;Cross-references: GB:AE000178; GB:U00096; NID:91786967; PMID:91786967; PMID:91786974;	
		A;Experimental source: strain K-12, substrate MG1655	
		C;Genes: galE	
		A;Gene: galE	
		A;Map position: 17 min	
		C;Function: isomerase	
		A;Description: isomerase	
		A;Pathway: galactose metabolism	
		C;Superfamily: Escherichia coli UDPglucose 4-epimerase	
		C;Keywords: galactose metabolism; isomerase	
		F1-336/Domain: UDPglucose 4-epimerase homology <UDP>	
		Query Match Score 49.7%; DB 4; Length 1259;	
Best Local Similarity 54.8%; Pred. No. 0.016;		Best Local Similarity 38.8%; Score 60.5%; DB 1; Length 338;	
Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;		Pred. No. 0.95%;	
Qy 2 APRFIQIQLDKREIDFNRLVVEFN-PUS 31		Mismatches 8; Indels 1;	
Db 120 APRFIQIQLDKREIDFNRLVVEFN-PUS 150			
RESULT 5			
B34C87	hypothetical protein (L1H 3' region) - human		
C;Species: Homo sapiens (man)			
C;Accession: B34C87			
R;Scott, A.F.; Schneckepfeppe, B.J.; Abdelraziq, M.; Comey, C.T.; O'Hara, B.; Rossiter, J.			
A;Title: Origin of the human L1 elements: proposed progenitor genes deduced from a conse			
A;Reference number: A34087; MUID:88085185; PMID:3692483			
A;Accession: B34087			
A;Status: preliminary; not compared with conceptual translation			
A;Molecule type: DNA			
A;Residues: 1-1280 <SCO>			
C;Superfamily: pol polyprotein			
Query Match Score 49.7%; DB 2; Length 1280;			
RESULT 7			
C90727	UDP-galactose-4-epimerase [imported] - Escherichia coli (strain O157:H7, substrain RIMD		

C; Species: Escherichia coli  
 C; Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
 C; Accession: C90727  
 R; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 R; Hayashi, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 R; Nagasawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A; Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
 A; Reference number: A96629; MUID:21156231; PMID:11258796  
 A; Accession: C90727  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-338 <HAY>  
 A; Cross-references: UNIPROT:Q8X942; UNIPROT:Q8FR9; GB:BAD00007; PID:91  
 A; Experimental source: strain O157:H7, substrate RIMD 0509952  
 C; Genetics:  
 A; Gene: ECs0787  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 Query Match 38 8% Score 60.5; DB 2; Length 338;  
 Best Local Similarity 46.2%; Pred. No. 0.95; Indels 1; Gaps 1;  
 Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;  
 Qy 6 IKQILDL-KREIDENVRLLVEYFNPL 30  
 Db 156 VEQILTDLQAZQPDMSIALRYFNPFV 181

RESULT 8  
 UDP-galactose 4-epimerase [imported] - Escherichia coli (strain O157:H7, substrate EDL93  
 C; Species: Escherichia coli  
 C; Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C; Accession: D85578  
 R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lin, A.; Dimalanta, E.; Potamitis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A; Reference number: A85480; MUID:21074935; PMID:1120551  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-338 <STO>  
 A; Cross-references: UNIPROT:Q8X942; UNIPROT:Q8FR9; GB:AE005174; PID:9125113888; PID:NID:AG  
 A; Experimental source: strain O157:H7, substrate EDL933  
 C; Genetics:  
 A; Gene: galE  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 Query Match 38 8% Score 60.5; DB 2; Length 338;  
 Best Local Similarity 46.2%; Pred. No. 0.95; Indels 1; Gaps 1;  
 Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;  
 Qy 6 IKQILDL-KREIDENVRLLVEYFNPL 30  
 Db 156 VEQILTDLQAZQPDMSIALRYFNPFV 181

RESULT 9  
 UDP-glucose 4-epimerase [imported] - Streptococcus pneumoniae (strain TIGR4)  
 C; Species: Streptococcus pneumoniae  
 C; Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 C; Accession: B95187  
 R; Reittelin, H.; Nelson, K.E.; Pauisen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,  
 nson, T.; Hickey, B.K.; Holt, I.B.  
 A; Authors: Lotrus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A; Reference number: A95000; MUID:21357209; PMID:1146616  
 A; Status: preliminary  
 A; Molecule type: DNA

A; Residues: 1-339 <KUR>  
 A; Cross-references: UNIPROT:Q9TPK2; GB:AE005672; PID:AAK75691.1; PID:914973098; GSPDB:G  
 A; Experimental source: strain TIGR4  
 C; Genes:  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 Query Match 37.5%; Score 58.5; DB 2; Length 339;  
 Best Local Similarity 38.7%; Pred. No. 1.8;  
 Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;  
 Qy 1 RAPRFKQILDL-KREIDENVRLLVEYFNPL 30  
 Db 152 RTKLMLBEILTDIYKADSEWWVLLYFNPI 182

RESULT 10  
 UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Streptococcus pneumoniae (strain R6)  
 C; Species: Streptococcus pneumoniae  
 C; Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C; Accession: C98054  
 R; Hoskins, J.A.; Alborn Jr., W.; Blasczak, L.; Burgett, S.; Dehoff, B.S.; E:  
 e, R.; LeBlanc, D.J.; Lee, L.; Lefkowitz, B.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc  
 Y, P.; Sun, P.M.; Winkler, M.E.  
 A; Residues: 1-339 <KUR>  
 A; Cross-references: UNIPROT:QBDNY6; GB:AE007317; PID:915459117; GSPDB:G  
 C; Genetics:  
 A; Gene: galE  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 C; Keywords: isomerase  
 A; Reference number: A97832; MUID:21429245; PMID:11544234  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-339 <STO>  
 A; Cross-references: UNIPROT:QBDNY6; GB:AE007317; PID:915459117; GSPDB:G  
 C; Genetics:  
 A; Gene: galE  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
 C; Keywords: isomerase  
 Query Match 37.5%; Score 58.5; DB 2; Length 339;  
 Best Local Similarity 38.7%; Pred. No. 1.8;  
 Matches 12; Conservative 9; Mismatches 9; Indels 1; Gaps 1;  
 Qy 1 RAPRFKQILDL-KREIDENVRLLVEYFNPL 30  
 Db 152 RTKLMLBEILTDIYKADSEWWVLLYFNPI 182

RESULT 11  
 UDPglucose 4-epimerase (EC 5.1.3.2) (clone GBP142) - guar  
 N; Alternate names: UDP-galactose 4-epimerase  
 C; Species: Cyamopsis tetragonoloba (Guar, cluster bean)  
 C; Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C; Accession: R10496  
 R; Brundt, J.; Joarsbo, M.; Pedersen, S.G.; Marcusen, J.  
 Plant Sci. 142, 147-154, 1999  
 A; Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase ext  
 A; Reference number: Z17038  
 A; Accession: R10496  
 A; Status: preliminary; translated from GB/ENBL/DDBJ  
 A; Molecule type: mRNA  
 A; Residues: 1-354 <BRU>  
 A; Cross-references: UNIPROT:Q65780; EMBL:AJ005081  
 C; Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase; UDP  
 C; Keywords: Galactose metabolism; isomerase  
 F; 12-350/Domain: UDPglucose 4-epimerase homology <UDP>  
 Query Match 36.9%; Score 57.5; DB 2; Length 354;  
 Best Local Similarity 32.3%; Pred. No. 2.6;  
 Matches 10; Conservative 12; Mismatches 8; Indels 1; Gaps 1;  
 Qy 1 RAPRFKQILDLKR-EIDFNRVRLVEYFNPL 30

RESULT 12

Qy S49879  
Db 163 RSKLFVEENWARDIQRAEAEMRILLRYFNFV 193

hypothetical carbonin-like 2S protein - fava bean

C;Species: Vicia faba (fava bean)

C;Accession: S49879  
C;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-285 <NON>  
A;Cross-references: UNIPROT:Q41660; EMBL:246826; NID:9600101; PID:9600102  
C;Superfamily: alcohol sulfotransferase

Query Match 36.5%; Score 57; DB 2; Length 285;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PRFIQILDLKREIDENVL 24  
Db 141 PRLIGOLITELKKERDLNTHVV 162

RESULT 13

Qy S49879  
Db T121257

nodulin - fava bean

C;Species: Vicia faba (fava bean)

C;Accession: T121257  
C;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-285 <PER>  
A;Cross-references: UNIPROT:Q41664; EMBL:246911; NID:91143616; PIDN:CAA87009\_1; PID:91143616  
A;Experimental source: strain Kleine Thuerlinger  
C;Superfamily: alcohol sulfotransferase

Query Match 36.5%; Score 57; DB 2; Length 285;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PRFIQILDLKREIDENVL 24  
Db 141 PRLIGOLITELKKERDLNTHVV 162

RESULT 14

Qy S49879  
Db T121257

hypothetical carbonin-like 2S protein (clone PVSNAG2) - spring vetch

C;Species: Vicia sativa (spring vetch, rare)

C;Accession: T121257  
C;Status: preliminary  
A;Description: A genomic sequence encoding putative narbonin from Vicia sativa.  
A;Reference number: S49880  
A;Molecule type: DNA  
A;Residues: 1-285 <NON>  
A;Cross-references: UNIPROT:Q41704; EMBL:246835; NID:9600112; PID:9600113



OX	NCBITaxID=9606;
RN	[1] _
SEQUENCE FROM N.A.	
MEDLINE=97285120; PubMed=9140393; Kimberland M.L., Naas T.P.,	
DR	Dombroski B.A., Moran J.V., Kazazian H.H. Jr.,
RX	DeBernardini R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.,
RA	"Many human L1 elements are capable of retrotransposition.";
DR	Nat. Genet. 16:37-43 (1997).
RT	
EMBL; U93563; AAC51261; 1;	
PIR; B28096;	B8096.
DR	
PIR; S23650;	S23650.
DR	GO: GO:000373; tRNA binding; IEA.
DR	GO: GO:0033964; tRNA-directed DNA polymerase activity; IEA.
DR	GO: GO:0006278; tRNA-dependent DNA replication; IEA.
DR	InterPro; IPR005135; Exo_endo_phos.
DR	InterPro; IPR000477; RVTse.
DR	
Pfam; PF00078; RVT.	1.
DR	
Pfam; PF00078; RVT.	1.
DR	tRNA-directed DNA polymerase; Transferase.
KW	RNA-directed DNA polymerase; Transferase.
SEQUENCE 1275 AA;	149062 MW;
SQ	350B4F0350E525FO CRC64;
RESULT 3	
000362	PRELIMINARY;
ID	000362
AC	000362;
DT	01-JUL-1997 (TREMBL); 04-Created)
DT	01-JUL-1997 (TREMBL); 04-Last sequence update)
DT	01-OCT-2003 (TREMBL); 25-Last annotation update)
DE	Purative p150.
OS	Homo sapiens (Human)
OC	Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBITaxID=9606;
RN	[1] _
SEQUENCE FROM N.A.	
MEDLINE=97285120; PubMed=9140393; Kimberland M.L., Naas T.P.,	
DR	Dombroski B.A., Moran J.V., Kazazian H.H. Jr.,
RX	DeBernardini R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.,
RA	"Many human L1 elements are capable of retrotransposition.";
DR	Nat. Genet. 16:37-43 (1997).
RT	
EMBL; U93564; AAC51263; 1;	
PIR; B28096;	B28096.
DR	
PIR; S23650;	S23650.
DR	GO: GO:000373; tRNA binding; IEA.
RA	GO: GO:0033964; tRNA-directed DNA polymerase activity; IEA.
DR	GO: GO:0006278; tRNA-dependent DNA replication; IEA.
DR	InterPro; IPR005135; Exo_endo_phos.
DR	InterPro; IPR000477; RVTse.
DR	
Pfam; PF00078; RVT.	1.

NCBI_TaxID=9606;		PRELIMINARY;	PRT;	1275 AA.
[1]				
RN	000363; 000363;	AC	AC	
SEQUENCE FROM N.A.		DT	DT	Created)
MEDLINE=97285120; PubMed=9140393;		DT	DT	Last sequence update)
RX	01-JUL-1997 (TREMBLrel. 04,			
RX	01-JUL-1997 (TREMBLrel. 04,			
Sassaman D.M., Domroeki B.A.; Moran J.V., Kimberland M.L., Naas T.P.,				
DeBarrardins R.J., Swergold G.D., Kazazian H.H., Jr.,				
DeBarrardins R.J., Gabriel A., Swergold G.D., Kazazian H.H., Jr.,				
"Many human Li elements are capable of retrotransposition.",				
RA	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
RA	DB			
RA	Putative P150.			
RA	OS			
RA	Homo sapiens (Human).			
RA	OC			
RA	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butteleostomi;			
RA	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RA	NCBI_TaxID=9606;			
PIR;	[1]	RN	RN	
PIR;	S23650; S23650.	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	
PIR;	S23650; S23650.	RX	RX	
GO;	GO:0003723; P:RNA binding; IEA.	Medline=97285120; PubMed=9140393;	Medline=97285120; PubMed=9140393;	
GO;	GO:0003964; P:RNA-directed DNA polymerase activity; IEA.	RA	RA	
GO;	GO:0006278; P:RNA-dependent DNA replication; IEA.	Sassaman D.M., Domroeki B.A., Moran J.V., Kimberland M.L., Naas T.P.	DeBarrardins R.J., Gabriel A., Swergold G.D., Kazazian H.H., Jr.,	
DR	DR	RA	RA	
DR	InterPro; IPR005135; Exo_endo_phos.	"Many human Li elements are capable of retrotransposition.";	"Many human Li elements are capable of retrotransposition.";	
DR	InterPro; IPR00477; RVTse.	Genet. 16:37-43 (1987).	Genet. 16:37-43 (1987).	
DR	Pfam; PF00372; Exo_endo_phos; 1.	RL	RL	
DR	Pfam; PF00078; RVTse; 1.	EMBL; AAC51264; 1; -.	EMBL; AAC51264; 1; -.	
DR	RNA-directed DNA Polymerase; Transferase.	DR	DR	
SEQUENCE	SEQUENCE 1275 AA; 149062 MW; 350BAF0358E525F0 CRC64;	PIR; B28096; B28096.	PIR; B28096; B28096.	
SEQUENCE		DR	DR	
DR	DR	PIR; S23650; S23650.	PIR; S23650; S23650.	
DR	DR	GO; GO:0003723; P:RNA binding; IEA.	GO; GO:0003723; P:RNA binding; IEA.	
DR	DR	GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.	GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.	
DR	DR	GO; GO:0006278; P:RNA-dependent DNA replication; IEA.	GO; GO:0006278; P:RNA-dependent DNA replication; IEA.	
DR	DR	IntePro; IPR005135; Exo_endo_phos.	IntePro; IPR005135; Exo_endo_phos.	
DR	DR	InterPro; IPR00477; RVTse.	InterPro; IPR00477; RVTse.	
DR	PFam; PF03172; Exo_endo_Phosph; 1.	DR	DR	
DR	PFam; PF00078; RVTse; 1.	PFam; PF00078; RVTse; 1.	PFam; PF00078; RVTse; 1.	
KW	RNA-directed DNA Polymerase; Transferase.	KW	KW	
SEQUENCE	SEQUENCE 1275 AA; 14907 MW; A866976EA3FD8F74 CRC64;	SEQUENCE 1275 AA; 14907 MW; A866976EA3FD8F74 CRC64;	SEQUENCE 1275 AA; 14907 MW; A866976EA3FD8F74 CRC64;	SEQUENCE 1275 AA; 14907 MW; A866976EA3FD8F74 CRC64;

```

RESULT 3
000362 PRELIMINARY; PRT; 1275 AA.
ID 000362; AC 000162; DT 01-JUL-1997 (TREMBL); 04, Created
AC 000162; DT 01-JUL-1997 (TREMBL); 04, Last sequence update)
DT 01-OCT-2003 (TREMBL); 25, Last annotation update)
DE Puratran p150
DE Homo sapiens (Human)
OS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9666;
OX
RN [1]
RN SEQUENCE FROM N.A.
RN MEDLINE=97285120; PubMed=9140393;
RN Sastre D.M., Dobroski B.A., Mcbran J.V., Kimberland M.L., Naas T.P.,
RN DeBarrardins R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.; ;
RN "Many human LINE elements are capable of retrotransposition.";
RN Natl. Genet. 16:32-43 (1997).
RN EMBL; US3564; AAC51263; 1. -
DR PIR; B28096; B28096.
PIR; S23650; S23650.
GO; GO:0003722; F:RNA binding; IEA.
GO; GO:0003946; F:RNA-directed DNA polymerase activity; IEA.
GO; GO:0006276; P:RNA-dependent DNA replication; IEA.
DR .InterPro; IPR05115; Exo_endo_phos.
DR .InterPro; IPR00477; Revise.
Pfam; PF03372; Exo_endo_phos; 1.
Pfam; PF00078; RVT; 1.
DR RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1275 AA; 149201 MW;

```

Query Match	Score 78.5;	DB 2;	Length 1275;
Best Local Similarity	50.3%;	Pred. No. 0.032;	
Matches 17;	Conservative	Mismatches 7;	Indels 1;
Qy	2 APRFIKQVQLDILREIDENFVRILVEYFN PLS 31		
Db	121 APRFIKQVQLSDLQRDLDSHTLINGDFNTPLS 151		
RESULT 5			
000366	PRELIMINARY;	PBT;	1275 AA.
AC	000366;		
DT	01-JUL-1997 (TREMBLrel. 04, Created)		
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Putative protein.		
OS	Homo sapiens (Human)		
OC	Metazoa; Chordata; Craniata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID	=9606;		
[1]			
RN	SEQUENCE FROM N.A.		
RX	MEDLINE=97285120; PubMed=9140393;		
RA	Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T., DeBarardins R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;		
PT	"Many human L1 elements are capable of retrotransposition."		
NAT	Genet. 16:37-43 (1997).		
DR	EMBL: U93567; AAC51267.1; -.		
DR	PIR: B28096; B28096.		
DR	PIR: S23550; S23560.		
DR	GO: GO:0003723; F: RNA binding; IEA.		
DR	GO: GO:0003944; F: RNA-directed DNA polymerase activity; IEA.		
DR	GO: GO:0006278; P: RNA-dependent DNA replication; IEA.		
DR	InterPro: IPR005135; Exo_endo_phos.		
DR	InterPro: IPR00477; Rvtse.		
PFAM	PF03372; Exo_endo_phos; 1.		
DR	PFam: PF00078; RVT; 1.		
KW	RNA-directed DNA polymerase; Transferase.		
SEQUENCE	1275 AA; 149033 MW; 07888F84DB831A2 CRC64;		
SQ	Query Match	Score 78.5;	DB 2;

Qy	2 APPRFIKQLLKLREIDFNVRLEYVN-PLS	31
Db	121 APPRFIKQVLSDLQRDLSHTLIMGDFNTPLS	151

Query	Match	Score	DB	Length
DR PFam; PF03312; Exo_endo_phos; 1.	PFam; PF00178; RVT; 1.	50.3 %	2	1275;
DR RNA-directed DNA Polymerase; Transferase	KW 149033 MW 07888F9F4DBB831A2 CRC64;			
SQ SEQUENCE 1275 AA;				

RESULT 4  
000363



RA	DeBarddinis R.J., Gabriel A., Swergold G.D., Kazazian H.H., Jr.; "Many human L1 elements are capable of retrotransposition.", Nat. Genet. 16:37-42 (1997).	DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
RL	EMBL; U93574; AAC51279.1; -.	DE Hypothetical protein
DR	PIR: B28096; B28096.	OC Homo sapiens (Human)
DR	PIR: JU0033; JU0033.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.
DR	NCBI_TaxID=9606;	OC NCBI_TaxID=9606;
DR	RN [1]	RN
DR	SEQUENCE FROM N.A.	RP
RC	TISSUE=Placenta;	RC
RA	Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanenori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,	RA
RA	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.	RA
DR	EMBL; AK131313; BAD18477.1; -.	DR
DR	InterPro; IPR00477; RVTse.	DR
DR	Pfam; PF01372; Exo_endo_phos; 1.	DR
DR	RNA-directed DNA Polymerase; Transferase.	DR
KW	SEQUENCE 1275 AA; 148879 MW; P00667P4BBBB72B87 CRC64;	DR
SQ	50.3%; Score 78.5; DB 2; Length 1275; Best Local Similarity 54.8%; Pred. No. 0.032; DB 2; Length 1275; Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;	DR
QY	2 APREFKQILDLKREIDENVRLVEYFN-PFLS 31	DR
Db	121 APRFIKQVLSDLQRDLSHFLIMEDNTPLS 151	DR
RESULT 10		DR
Q8TE30	PRELIMINARY; PRT; 1275 AA.	DR
ID	Q8TE30	DR
AC	Q8TE30_2002 (TrEMBLrel. 21, Created)	DR
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	DT
DT	01-OCT-2002 (TrEMBLrel. 25, Last annotation update)	DT
DE	Hypothetical protein.	DE
OS	Homo sapiens (Human).	OS
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrini; Hominidae; Homo.	OC
OC	SEQUENCE FROM N.A.	OC
RP	MEDLINE=166838; PubMed=11810275;	RP
RX	Berries S.M., Morris C.M.; "A full-length and potentially active LINE element is integrated polymorphically within the IGL locus in a genetically unstable region of chromosome 22."	RX
RT	"A full-length and potentially active LINE element is integrated polymorphically within the IGL locus in a genetically unstable region of chromosome 22."	RT
RT	RT	RT
RJ	Hum. Genet. 109:628-637(2001).	RJ
DR	PIR: B28096; B28096.	DR
DR	PIR: S23550; S23650.	DR
DR	GO; GO:0003964; RNA-directed DNA polymerase activity; IEA.	DR
DR	GO; GO:0006278; RNA-dependent DNA replication; IEA.	DR
DR	InterPro; IPR005135; Exo_endo_phos.	DR
DR	IntePro; IPR00477; RVTse.	DR
DR	Pfam; PF000728; RVTse.	DR
XW	Hypothetical protein; RNA-directed DNA polymerase; Transferase.	XW
SEQUENCE 1275 AA; 149009 MW; B3279D50581764 CRC64;	SEQUENCE 1275 AA; 149009 MW; B3279D50581764 CRC64;	
QY	2 APREFKQILDLKREIDENVRLVEYFN-PFLS 31	QY
Db	121 APRFIKQVLSDLQRDLSHFLIMEDNTPLS 151	Db
RESULT 11		DR
Q6ZN98	PRELIMINARY; PRT; 244 AA.	DR
ID	Q6ZN98	DR
AC	Q6ZN98; Q6ZN98.	AC
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	DT

AC P08547; DT 01-AUG-1988 (Rel. 08, Created) DT 01-AUG-1988 (Rel. 08, Last sequence update) DT 05-JUL-2004 (Rel. 44, Last annotation update) LINE-1 reverse transcriptase homolog.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI TaxID=9606;

RN [1]

SEQUENCE.

RX MEDLINE#6230917; PubMed=2423883;

RA Hattori M., Kuhara S., Takezawa O., Sakaki Y.;

RT "L1 family of repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transciptase-related protein.";

RL Nature 321:625-628(1986).

CC -!- MISCELLANEOUS: This sequence was constructed from an alignment of published and unpublished sequences, determined in various laboratories, belonging to the LINE-1 family.

DR A25213; GNHJ11.

CC InterPro; IPR005135; Exo\_endo\_phos.

DR InterPro; IPR000477; RVTSe.

DR Pfam; PF03372; Exo\_endo\_phos; 1.

DR Pfam; PF00078; RTT; 1.

DR Direct protein sequencing; RNA-directed DNA polymerase.

SQ SEQUENCE 1259 AA; 14712 MW; F3BC18A38039195 CRC64;

Query Match 49.7%; Score 77.5; DB 1; Length 1259;

Best Local Similarity 54.8%; Pred. No. 0.044; Matches 17; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

Db 120 APRFIQVILDLKREIDENYRVLVEYFN-PLS 31

RESULT 14

Q6ZNC3 PRELIMINARY; PRT; 314 AA.

AC Q6ZNC3;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Hypothetical protein FLJ16220.

OS Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tongue;

RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S., Watanabe M., Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M., Sugiyama T., Irie R., Otuki T., Saito K., Watanabe A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Matsuo Y., Nagai K., Isogai T.;

DR Submitted (MAR 2004) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK131275; BAD18452.1; -.

FT NON\_TER 314

SQ SEQUENCE 314 AA; 36311 MW; 2C37DD7DD1001A56 CRC64;

Query Match 48.4%; Score 75.5; DB 2; Length 314;

Best Local Similarity 51.6%; Pred. No. 0.019; Matches 16; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

Db 121 APRFIQVILDLKREIDENYRVLVEYFN-PLS 31

RESULT 15

BAD18452 PRELIMINARY; PRT;

AC BAD18452;

DT 12-MAY-2004 (TREMBLrel. 27, Last sequence update)

DT 12-MAY-2004 (TREMBLrel. 27, Last annotation update)

DE CDNA FLJ16220 firs, clone CTONG3002552. (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Tongue;

RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S., Watanabe M., Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M., Sugiyama T., Irie R., Otuki T., Saito K., Nishikawa T., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Matsuo Y., Nagai K., Isogai T.;

DR "NEDO human cDNA sequencing project.";

RL Submitted (MAR 2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK131275; BAD18452.1; -.

FT NON\_TER 314

SQ SEQUENCE 314 AA; 36311 MW; 2C37DD7DD1001A56 CRC64;

Query Match 48.4%; Score 75.5; DB 2; Length 314;

Best Local Similarity 51.6%; Pred. No. 0.019; Matches 16; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

Db 121 APRFIQVILDLKREIDENYRVLVEYFN-PLS 151

Search completed: November 10, 2004, 13:38:53

Job time : 39.8532 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: November 10, 2004, 13:44:14 ; Search time 9.51698 Seconds  
(without alignments)

Title: US-10-092-750-44

Perfect score: 125

Sequence: 1 IVAIAGRLRMLQFNGELEASAKN 26

Scoring table: BL0SM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cgcn2\_6/pododata/1/1aa/'5A COMB.pep'\*  
 2: /cgcn2\_6/pododata/1/1aa/'5B COMB.pep'\*  
 3: /cgcn2\_6/pododata/1/1aa/'6A COMB.pep'\*  
 4: /cgcn2\_6/pododata/1/1aa/'6B COMB.pep'\*  
 5: /cgcn2\_6/pododata/1/1aa/'PETUS COMB.pep'\*  
 6: /cgcn2\_6/pododata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.6	3519	3	US-09-428-517-4
2	49	39.2	3816	3	US-09-428-517-3
3	49	39.2	4150	3	US-09-428-517-2
4	44.5	35.6	40	4	US-09-430-915A-82
5	44.5	35.6	40	4	US-09-430-917-82
6	44.5	35.6	40	4	US-09-434-000C-3390
7	44	35.2	268	4	US-09-434-000C-5975
8	44	35.2	315	4	US-09-434-000C-27945
9	44	35.2	415	4	US-09-432-911A-33751
10	43.5	34.8	48	4	US-09-433-005A-16
11	43.5	34.8	54	1	US-08-167-035-28
12	43.5	34.8	54	1	US-08-108-287-28
13	43.5	34.8	54	2	US-08-139-005-28
14	43.5	34.8	54	3	US-09-246-510B-19
15	43.5	34.8	54	3	US-09-380-598-30
16	43.5	34.8	348	4	US-09-107-532A-5421
17	43.5	34.8	800	6	5183745-3
18	43.5	34.8	1130	4	US-09-518-032-834
19	43	34.4	1025	2	US-08-330-792D-23
20	43	34.4	1026	2	US-08-330-792D-22
21	42.5	34.0	271	4	US-09-255-991A-28051
22	42	33.6	16	4	US-09-514-664B-46
23	42	33.6	27	4	US-09-444-664B-19
24	42	33.6	161	4	US-10-104-664A-729
25	42	33.6	349	4	US-09-252-991A-23984
26	42	33.6	502	4	US-09-252-991A-27619
27	42	33.6	530	3	US-09-199-637A-120

## ALIGNMENTS



```

CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION DOCUMENT NUMBER: 18 872
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-09-879-957-82

Query Match 35.6%; Score 44.5%; DB 4; Length 40;
Best Local Similarity 57.9%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 5;
Indels 1; Gaps 1;

Qy 8 RLRLMGDDQENGELEASAKN 26
Db 21 KLRVGYNNNG-EAQTKN 38

RESULT 6
US-09-134-000C-5390
Sequence 5390, Application US/09134000C
; Pat. No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C.
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn, version 3.1
SEQ ID NO 5390
LENGTH: 351
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-5390

Query Match 35.6%; Score 44.5%; DB 4; Length 351;
Best Local Similarity 40.7%; Pred. No. 22;
Matches 11; Conservative 6; Mismatches 7;
Indels 3; Gaps 1;

Qy 1 IVAAIGRLRMLGDDQENGELEASA 24
Db 258 IVAAGTTSIRIETGTKFNGEIQADS 284

RESULT 7
US-09-134-000C-5975
Sequence 5975, Application US/09134000C
; Pat. No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
; Pat. No. 6617156
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfeld et al
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196-136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-01-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27945
LENGTH: 315
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27945

Query Match 35.2%; Score 44%; DB 4; Length 315;
Best Local Similarity 45.8%; Pred. No. 23;
Matches 11; Conservative 4; Mismatches 9;
Indels 0; Gaps 0;

Qy 3 ALIAGTRMLGDDQENGELEASAKN 26
Db 285 ATVAGRLAYLAGMRPLKLYASSS 308

RESULT 9
US-09-252-991A-23751
Sequence 23751, Application US/09252991A
; Pat. No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfeld et al
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196-136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23751
LENGTH: 415
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23751

Query Match 35.2%; Score 44%; DB 4; Length 415;
Best Local Similarity 43.5%; Pred. No. 33;
```

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0; REFERENCE/DOCKET NUMBER: 7683-062

Query 2 VAIAGGRLRLGDDQFNGELEASA 24 ; TELECOMMUNICATION INFORMATION:

DB 142 VALLRGRQRPLHDHFGGRDRHRA 164 ; TELEPHONE: (212) 790-9690

RESULT 10 ; TELEFAX: (212) 869-9741/8864

US-023-905A-16 ; TELEX: 66141 PENNIE

/ Sequence 16, Application US/09023905A ; INFORMATION FOR SEQ ID NO: 28:

GENERAL INFORMATION: ; SEQUENCE CHARACTERISTICS:

/ APPLICANT: Roberts, Thomas M. ; LENGTH: 54 amino acids

/ PATENT NO.: 6415778 ; TYPE: amino acid

/ ATTORNEY/AGENT INFORMATION: ; TOPOLOGY: unknown

/ APPLICANT: King, Frederick J. ; MOLECULE TYPE: protein

/ APPLICANT: Harris, David F. ; US-08-167-035-28

/ APPLICANT: Hu, Erding ;

/ APPLICANT: Spiegelman, Bruce ;

/ APPLICANT: Chan, Joanne ;

/ TITLE OF INVENTION: Differentiation Enhancing Factors and Uses ;

/ FILE REFERENCE: DFN-021 ; CURRENT APPLICATION NUMBER: US/09/023, 905A

/ CURRENT FILING DATE: 1998-02-13 ; PRIORITY APPLICATION NUMBER: US 60/038, 191

/ PRIOR FILING DATE: 1997-02-14 ; NUMBER SEQ ID NOS: 39

/ SOFTWARE: FastSEQ for Windows Version 4.0 ;

/ SEQ ID NO 16 ; LENGTH: 48

/ TYPE: PRT ; ORGANISM: Mus musculus

US-09-023-905A-16 ; ADDRESS: PENNIE & EDMONDS

Query Match 34.8%; Score 43.5%; DB 4; Length 48; COMPUTER READABLE FORM:

Best Local Similarity 55.0%; Pred. No. 2.9; Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1; MEDIUM TYPE: Floppy disk

Query 8 RLRLMGDQFNGE-LEASAKN 26 ; COMPUTER: IBM PC compatible

DB 20 KLRVLYNHNGEWCQAQTKN 39 ; OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 11 ; SOFTWARE: PatientIn Release #1.0, Version #1.3.0

US-08-167-035-28 ; ATTORNEY/AGENT INFORMATION:

/ Sequence 28, Application US/08167035 ; CURRENT APPLICATION DATA:

/ PATENT NO.: 567421 ; APPLICATION NUMBER: US/09/208, 887A

/ GENERAL INFORMATION: ; CURRENT APPLICATION DATA:

/ APPLICANT: Schlessinger, Joseph ; APPLICATION NUMBER: US/09/208, 887A

/ APPLICANT: Skolnick, Edward Y. ; FILING DATE: 11-MAR-1994

/ APPLICANT: Margolis, Benjamin L. ; CLASSIFICATION: 435

/ TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR ; ATTORNEY/AGENT INFORMATION:

/ TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE ; NAME: Coruzzi, Laura A.

/ TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS ; REGISTRATION NUMBER: 30,742

/ NUMBER OF SEQUENCES: 51 ; REFERENCE/DOCKET NUMBER: 7683-063

/ CORRESPONDENCE ADDRESS: ZIP: 10036-2711

/ STREET: 1155 Avenue of the Americas ; TELECOMMUNICATION INFORMATION:

/ CITY: New York ; TELEPHONE: (212) 790-9090

/ STATE: New York ; TELEFAX: (212) 869-9741/8864

/ COUNTRY: ZIP: 10036-2711 ; TELEX: 66141 PENNIE

RESULT 12 ; INFORMATION FOR SEQ ID NO: 28:

US-08-208-887A-28 ; SEQUENCE CHARACTERISTICS:

/ Query Match 34.8%; Score 43.5%; DB 1; Length 54;

/ Best Local Similarity 55.0%; Pred. No. 3.4; Matches 11; Conservation 2; Mismatches 6; Indels 1; Gaps 1;

/ MOLECULE TYPE: protein

Query 8 RLRLMGDQFNGE-LEASAKN 26 ;

DB 20 KLRVLYNHNGEWCQAQTKN 39 ;

Query 8 RLRLMGDQFNGE-LEASAKN 26 ;

DB 20 KLRVLYNHNGEWCQAQTKN 39 ;

RESULT 13  
US-08-539-005-28  
Sequence 28, Application US/08539005  
GENERAL INFORMATION:  
Patent No. 5,858,686  
APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnick, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR EUKARYOTIC TYROSINE KINASES AND NOVEL TARGET PROTEINS  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESS: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: US  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DCS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,005  
FILING DATE: 4-OCT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,035  
FILING DATE: 16-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 869-9090  
TELEFAX: (212) 869-9741/3864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 54 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-539-005-28

Query Match 34.8%; Score 43.5; DB 3; Length 54;  
Best Local Similarity 55.0%; Pred. No. 3.4;  
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
Qy 8 RLRLMGDQFNGE-LEASAKN 26  
Db 20 KLRLVLYNHNGEWCQAQTKN 39

RESULT 14  
US-09-346-510B-19  
Sequence 19, Application US/09346510B  
GENERAL INFORMATION:  
Patent No. 6,281,014  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Wang, Yinxiang  
TITLE OF INVENTION: SH3-Containing Protein, DNA and Uses Thereof  
FILE REFERENCE: D5221CIP  
CURRENT APPLICATION NUMBER: US/09/346,510B  
CURRENT FILING DATE: 1999-07-01  
PRIOR APPLICATION NUMBER: 08/671,732  
PRIOR FILING DATE: 1997-06-09  
NUMBER OF SEQ ID NOS: 32

Query Match 34.8%; Score 43.5; DB 3; Length 54;  
Best Local Similarity 55.0%; Pred. No. 3.4;  
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
Qy 8 RLRLMGDQFNGE-LEASAKN 26  
Db 20 KLRLVLYNHNGEWCQAQTKN 39

RESULT 15  
US-09-280-598-30  
Sequence 30, Application US/09280598  
GENERAL INFORMATION:  
Patent No. 6,391,584  
APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnick, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
APPLICANT: APP, Harold A.  
APPLICANT: TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND NOVEL TARGET PROTEINS  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,598  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-067  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 791-9090  
TELEFAX: (212) 869-9741/3864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 54 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-280-598-30

Query Match 34.8%; Score 43.5; DB 3; Length 54;  
Best Local Similarity 55.0%; Pred. No. 3.4;  
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
Qy 8 RLRLMGDQFNGE-LEASAKN 26  
Db 20 KLRLVLYNHNGEWCQAQTKN 39

Fri Nov 12 14:55:27 2004

us-10-092-750-44.ra1

Page 6

Search completed: November 10, 2004, 14:55:33  
Job time : 9.56698 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - Protein search, using SW mode!

Run on: November 11, 2004, 01:28:30 ; Search time 30.0717 Seconds  
 (without alignments) ; 305.399 Million cell updates/sec

Title: US-10-092-750-44

Perfect score: 125

Sequence: 1 IVAIAGRLRMLGDQENGELBAAKN 26

Scoring table: BUOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
 1: /cgn2\_6/podata/1/pubpaas/us00\_PUBCOMB.pep:  
 2: /cgn2\_6/podata/1/pubpaas/pct\_new\_pub.pep:  
 3: /cgn2\_6/podata/1/pubpaas/us06\_PUBCOMB.pep:  
 4: /cgn2\_6/podata/1/pubpaas/us07\_PUBCOMB.pep:  
 5: /cgn2\_6/podata/1/pubpaas/us07\_PUBCOMB.pep:  
 6: /cgn2\_6/podata/1/pubpaas/pctus\_pubcomb.pep:  
 7: /cgn2\_6/podata/1/pubpaas/us08\_PUBCOMB.pep:  
 8: /cgn2\_6/podata/1/pubpaas/us08\_PUBCOMB.pep:  
 9: /cgn2\_6/podata/1/pubpaas/us09a\_PUBCOMB.pep:  
 10: /cgn2\_6/podata/1/pubpaas/us09b\_PUBCOMB.pep:  
 11: /cgn2\_6/podata/1/pubpaas/us09c\_PUBCOMB.pep:  
 12: /cgn2\_6/podata/1/pubpaas/us09\_new\_pub.pep:  
 13: /cgn2\_6/podata/1/pubpaas/us10\_PUBCOMB.pep:  
 14: /cgn2\_6/podata/1/pubpaas/us10b\_PUBCOMB.pep:  
 15: /cgn2\_6/podata/1/pubpaas/us10c\_PUBCOMB.pep:  
 16: /cgn2\_6/podata/1/pubpaas/us10d\_PUBCOMB.pep:  
 17: /cgn2\_6/podata/1/pubpaas/us10\_new\_pub.pep:  
 18: /cgn2\_6/podata/1/pubpaas/us11a\_new\_pub.pep:  
 19: /cgn2\_6/podata/1/pubpaas/us11b\_new\_pub.pep:  
 20: /cgn2\_6/podata/1/pubpaas/us60\_PUBCOMB.pep:  
 RESULT 1 US-10-092-750-44

; Sequence 44, Application US-10092750  
 ; Publication No. US0030032157A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hammond, Philip W.  
 ; APPLICANT: Alpin, Julia  
 ; APPLICANT: Wright, Martin C.  
 ; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL  
 ; FILE REFERENCE: 50036/050002  
 ; CURRENT APPLICATION NUMBER: US-10/092,750  
 ; PRIORITY APPLICATION NUMBER: US 2001-03-08  
 ; NUMBER OF SEQ ID NOS: 253  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO 44  
 ; LENGTH: 26  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-092-750-44

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	26	14 US-10-092-750-44	Sequence 44, App1
2	88	70.4	71	17 US-10-092-750-44	Sequence 238718,
3	52	41.6	363	9 US-09-925-300-1530	Sequence 1530, Ap
4	52	41.6	1072	15 US-10-307-817-90	Sequence 90, App1
5	52	41.6	1072	15 US-10-287-226-230	Sequence 230, App
6	52	41.6	1091	15 US-10-307-817-469	Sequence 469, App
7	52	41.6	1091	15 US-10-287-226-659	Sequence 669, App
8	52	41.6	1095	15 US-10-307-817-86	Sequence 86, App
9	52	41.6	1095	15 US-10-287-226-232	Sequence 232, App
10	52	41.6	1095	15 US-10-287-226-238	Sequence 238, App
11	52	41.6	1095	15 US-10-287-226-446	Sequence 668, App
12	52	41.6	1100	15 US-10-287-226-668	Sequence 84, App
13	52	41.6	1101	15 US-10-307-817-84	

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	26	14 US-10-092-750-44	Sequence 44, App1
2	88	70.4	71	17 US-10-092-750-44	Sequence 238718,
3	52	41.6	363	9 US-09-925-300-1530	Sequence 1530, Ap
4	52	41.6	1072	15 US-10-307-817-90	Sequence 90, App1
5	52	41.6	1072	15 US-10-287-226-230	Sequence 230, App
6	52	41.6	1091	15 US-10-307-817-469	Sequence 469, App
7	52	41.6	1091	15 US-10-287-226-659	Sequence 669, App
8	52	41.6	1095	15 US-10-307-817-86	Sequence 86, App
9	52	41.6	1095	15 US-10-287-226-232	Sequence 232, App
10	52	41.6	1095	15 US-10-287-226-238	Sequence 238, App
11	52	41.6	1095	15 US-10-287-226-446	Sequence 668, App
12	52	41.6	1100	15 US-10-287-226-668	Sequence 84, App
13	52	41.6	1101	15 US-10-307-817-84	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy	Db	Match	Local Similarity	Mismatches	Indels	Gaps	Length
Qy	Db	1 IVAIAGRLRMLGDQNGELEASAKN 26	1 IVAIAGRLRMLGDQNGELEASAKN 26	0	0	0	26

RESULTS

Qy
----

; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yina  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 3-21(5322)B  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326.  
 ; SEQ ID NO 238718  
 ; LENGTH: 71  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_14929C.1.pep  
 ; US-10-425-115-238718  
 ;  
 Query Match 2 VAIAGCRRLMLQDFENGEL 20  
 Best Local Similarity 70.4%; Score 88; DB 17; Length 71;  
 Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 53 VAIVAGRRLMLQDFENGEL 1  
  
 RESULT 3  
 US-09-925-300-1530  
 ; Sequence 1530, Application US/09925300  
 ; Patent No. US200215161A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Rubin,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1530  
 ; LENGTH: 363  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (178)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (179)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-925-300-1530  
 ;  
 Query Match 1 IVAIIAGRRLMLQDFENGEL 25  
 Best Local Similarity 40.0%; Score 52; DB 9; Length 363;  
 Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
 Db 182 LVSSLTSGLLTIGDRFGALDAAK 206  
  
 RESULT 4  
 US-10-307-817-90  
 ; Sequence 90, Application US/10307817  
 ; Publication No. US20040058338A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agee et al.  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-502C  
 ; CURRENT FILING DATE: 2002-10-307,817  
 ; NUMBER OF SEQ ID NOS: 682  
 ; SEQ ID NO 90  
 ; LENGTH: 1072  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-307-817-90  
  
 Query Match 1 IVAIIAGRRLMLQDFENGEL 25  
 Best Local Similarity 40.0%; Score 52; DB 15; Length 1072;  
 Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
 Db 891 LVSSLTSGLLTIGDRFGALDAAK 915  
  
 RESULT 5  
 US-10-287-226-230  
 ; Sequence 230, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agee, Michele L., Alsobrook, John P.,  
 ; APPLICANT: Berdige, Constance, Boldge, Florence,  
 ; APPLICANT: Burgess, Catherine E., Chant, John S.,  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A., Edinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew, Ellerman, Karen, Ganguli, Esha A.,  
 ; APPLICANT: Gorman, Linda, Gerlach, Valerie, Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh, Khramtsov, Nikolai, Li, Li,  
 ; APPLICANT: Mezies, Peter S., Miller, Charles E., Milet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng, Ort, Tatiana,  
 ; APPLICANT: Padigaru, Muralidhara, Patterson, Meera, Rastelli, Luca,  
 ; APPLICANT: Rieger, Daniel K., Rothenberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G., Spaderna, Steven K.,  
 ; APPLICANT: Spytek, Kimberly A., Taupier, Jr., Raymond J.,  
 ; APPLICANT: Vernet, Corine A.M., Zernusen, Bryan D.,  
 ; APPLICANT: Zheng, Mei  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-480C  
 ; CURRENT FILING DATE: 2002-11-04  
 ; PRIORITY: 2002-02-04  
 ; PRIORITY NUMBER: 60/334,421  
 ; PRIORITY FILING DATE: 2001-11-30  
 ; PRIORITY NUMBER: 60/354,392  
 ; PRIORITY FILING DATE: 2002-02-04  
 ; PRIORITY NUMBER: 60/360,148  
 ; PRIORITY FILING DATE: 2002-02-27  
 ; PRIORITY NUMBER: 60/364,000

PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/404,821  
 PRIOR FILING DATE: 2002-08-20  
 PRIOR APPLICATION NUMBER: 60/334,526  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,409  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/354,227  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/334,027  
 PRIOR FILING DATE: 2001-11-28  
 PRIOR APPLICATION NUMBER: 60/311,641  
 PRIOR FILING DATE: 2001-11-20  
 REMAINING Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 673  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO: 230  
 LENGTH: 1072  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-287-226-230

Query Match 41.6%; Score 52; DB 15; Length 1072;  
 Best Local Similarity 40.0%; Pred. No. 24; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IVAIAGRLMLGDOFNGLEASAK 25  
 Db 891 LVSSITSGLTLTIGDRGALDAAK 915

---

RESULT 6  
 US-10-307-817-469  
 / Sequence 469, Application US/10307817  
 / Publication No. US20040058338A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Agee et al.  
 / TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 / CURRENT APPLICATION NUMBER: US/10/307,817  
 / CURRENT FILING DATE: 2002-12-02  
 / SOFTWARE: CuraseqList version 0.1  
 / SEQ ID NO: 469  
 / LENGTH: 1091  
 / TYPE: PRT  
 / ORGANISM: Mus musculus  
 US-10-307-817-469

Query Match 41.6%; Score 52; DB 15; Length 1091;  
 Best Local Similarity 40.0%; Pred. No. 24; Mismatches 7; Indels 0; Gaps 0;

Qy 1 IVAIAGRLMLGDOFNGLEASAK 25  
 Db 910 LVSSITSGLTLTIGDRGALDAAK 934

---

RESULT 7  
 US-10-287-226-669  
 / Sequence 669, Application US/10287226  
 / Publication No. US20040086875A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Agee, Michele L.,  
 / APPLICANT: Alsobrook, John P.,  
 / APPLICANT: Berbs, Constance,  
 / APPLICANT: Boldog, Ference,  
 / APPLICANT: Burgess, Catherine E.,  
 / APPLICANT: Chant, John S.,  
 / APPLICANT: Chaudhuri, Amitabha,  
 / APPLICANT: DiPippo, Vincent A.,  
 / APPLICANT: Edinger, Schonit R.,  
 / APPLICANT: Eisen, Andrew,

; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Ganguli, Ishita A.,  
 ; APPLICANT: Gorman, Linda,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramtsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malanyankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatiana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastelli, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothemberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,  
 ; APPLICANT: Spytek, Kimberley A.,  
 ; APPLICANT: Taupier, Jr., Raymond J.,  
 ; APPLICANT: Vernet, Corine A.M.,  
 ; APPLICANT: Zerhusen, Bryan D.,  
 / TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 / FILE REFERENCE: 21402-480C  
 / CURRENT APPLICATION NUMBER: US/10/287,226  
 / CURRENT FILING DATE: 2002-11-04  
 / PRIOR APPLICATION NUMBER: 60/334,421  
 / PRIOR FILING DATE: 2001-11-30  
 / PRIOR APPLICATION NUMBER: 60/354,392  
 / PRIOR FILING DATE: 2002-02-04  
 / PRIOR APPLICATION NUMBER: 60/360,148  
 / PRIOR FILING DATE: 2002-02-27  
 / PRIOR APPLICATION NUMBER: 60/364,000  
 / PRIOR FILING DATE: 2002-03-13  
 / PRIOR APPLICATION NUMBER: 60/404,821  
 / PRIOR FILING DATE: 2002-08-20  
 / PRIOR APPLICATION NUMBER: 60/334,526  
 / PRIOR FILING DATE: 2001-11-30  
 / PRIOR APPLICATION NUMBER: 60/354,409  
 / PRIOR FILING DATE: 2002-02-04  
 / PRIOR APPLICATION NUMBER: 60/364,227  
 / PRIOR FILING DATE: 2002-03-13  
 / PRIOR APPLICATION NUMBER: 60/334,027  
 / PRIOR FILING DATE: 2001-11-28  
 / PRIOR APPLICATION NUMBER: 60/331,641  
 / PRIOR FILING DATE: 2001-11-20  
 / Remaining Prior Application data removed - See File Wrapper or PALM.  
 / NUMBER OF SEQ ID NOS: 673  
 / SOFTWARE: CuraseqList version 0.1  
 / SEQ ID NO: 669  
 / LENGTH: 1091  
 / TYPE: PRT  
 / ORGANISM: Mus musculus  
 US-10-287-226-669

Query Match 41.6%; Score 52; DB 15; Length 1091;  
 Best Local Similarity 40.0%; Pred. No. 24; Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIAGRLMLGDOFNGLEASAK 25  
 Db 910 LVSSITSGLTLTIGDRGALDAAK 934

RESULT 8  
 US-10-287-226-86  
 / Sequence 86, Application US/10307817  
 / Publication No. US20040058338A1

GENERAL INFORMATION:  
 APPLICANT: Agee et al.  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-502C  
 CURRENT APPLICATION NUMBER: US/10/307,817  
 CURRENT FILING DATE: 2002-12-02  
 NUMBER OF SEQ ID NOS: 682  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO: 86  
 LENGTH: 1995  
 ORGANISM: Homo sapiens  
 -10-307-817-86

RESULT 9  
 -10-287-226-232  
 Sequence 232, Application US/10287226  
 Publication No. US20040086875A1  
 GENERAL INFORMATION:  
 APPLICANT: Agee, Michele L.,  
 APPLICANT: Alsobrook, John P.,  
 APPLICANT: Bergs, Constance,  
 APPLICANT: Bolod, Ference,  
 APPLICANT: Burgess, Catherine E.,  
 APPLICANT: Chaut, John S.,  
 APPLICANT: Chaudhuri, Amitabha,  
 APPLICANT: DiPippo, Vincent A.,  
 APPLICANT: Edinger, Andrew R.,  
 APPLICANT: Eisen, Andrew,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Ganguli, Esha A.,  
 APPLICANT: Gorman, Linda,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Ji, WeiZhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Khramtsov, Nikolai,  
 APPLICANT: Li, Li,  
 APPLICANT: Malyankar, Uriel M.,  
 APPLICANT: MacDougal, John R.,  
 APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Patterson, Meera,  
 APPLICANT: Rastelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark E.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spaderna, Steven K.,  
 APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taupier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zerhusen, Bryan D.,  
 APPLICANT: Zheng, Mei

FILE REFERENCE: 21402-480C  
 CURRENT APPLICATION NUMBER: US/10/287,226  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2000-11-30  
 PRIOR APPLICATION NUMBER: 60/354,392  
 PRIOR FILING DATE: 2002-02-04

RESULT 10  
 -10-287-226-238  
 Sequence 238, Application US/10287226  
 Publication No. US20040086875A1  
 GENERAL INFORMATION:  
 APPLICANT: Agee, Michele L.,  
 APPLICANT: Alsobrook, John P.,  
 APPLICANT: Bergs, Constance,  
 APPLICANT: Bolod, Ference,  
 APPLICANT: Burgess, Catherine E.,  
 APPLICANT: Chant, John S.,  
 APPLICANT: Chaudhuri, Amitabha,  
 APPLICANT: DiPippo, Vincent A.,  
 APPLICANT: Edinger, Andrew R.,  
 APPLICANT: Eisen, Andrew,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Ganguli, Esha A.,  
 APPLICANT: Gorman, Linda,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Ji, WeiZhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Khramtsov, Nikolai,  
 APPLICANT: Li, Li,  
 APPLICANT: Malyankar, Uriel M.,  
 APPLICANT: MacDougal, John R.,  
 APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Patterson, Meera,  
 APPLICANT: Rastelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark E.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spaderna, Steven K.,  
 APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taupier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zerhusen, Bryan D.,  
 APPLICANT: Zheng, Mei

FILE REFERENCE: 21402-480C  
 CURRENT APPLICATION NUMBER: US/10/287,226  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2000-11-30  
 PRIOR APPLICATION NUMBER: 60/354,392  
 PRIOR FILING DATE: 2002-02-04

RESULT 11  
 -10-287-226-238  
 Sequence 238, Application US/10287226  
 Publication No. US20040086875A1  
 GENERAL INFORMATION:  
 APPLICANT: Agee, Michele L.,  
 APPLICANT: Alsobrook, John P.,  
 APPLICANT: Bergs, Constance,  
 APPLICANT: Bolod, Ference,  
 APPLICANT: Burgess, Catherine E.,  
 APPLICANT: Chant, John S.,  
 APPLICANT: Chaudhuri, Amitabha,  
 APPLICANT: DiPippo, Vincent A.,  
 APPLICANT: Edinger, Andrew R.,  
 APPLICANT: Eisen, Andrew,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Ganguli, Esha A.,  
 APPLICANT: Gorman, Linda,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Ji, WeiZhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Khramtsov, Nikolai,  
 APPLICANT: Li, Li,  
 APPLICANT: Malyankar, Uriel M.,  
 APPLICANT: MacDougal, John R.,  
 APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Patterson, Meera,  
 APPLICANT: Rastelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark E.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spaderna, Steven K.,  
 APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taupier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zerhusen, Bryan D.,  
 APPLICANT: Zheng, Mei

RESULT 12  
 -10-287-226-238  
 Sequence 238, Application US/10287226  
 Publication No. US20040086875A1  
 GENERAL INFORMATION:  
 APPLICANT: Agee, Michele L.,  
 APPLICANT: Alsobrook, John P.,  
 APPLICANT: Bergs, Constance,  
 APPLICANT: Bolod, Ference,  
 APPLICANT: Burgess, Catherine E.,  
 APPLICANT: Chant, John S.,  
 APPLICANT: Chaudhuri, Amitabha,  
 APPLICANT: DiPippo, Vincent A.,  
 APPLICANT: Edinger, Andrew R.,  
 APPLICANT: Eisen, Andrew,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Ganguli, Esha A.,  
 APPLICANT: Gorman, Linda,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Ji, WeiZhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Khramtsov, Nikolai,  
 APPLICANT: Li, Li,  
 APPLICANT: Malyankar, Uriel M.,  
 APPLICANT: MacDougal, John R.,  
 APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Patterson, Meera,  
 APPLICANT: Rastelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark E.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spaderna, Steven K.,  
 APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taupier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zerhusen, Bryan D.,  
 APPLICANT: Zheng, Mei

APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zerhusen, Bryan D.,  
 APPLICANT: Zhong, Mei  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 CURRENT APPLICATION NUMBER: US/10/287,226  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,392  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/360,148  
 PRIOR FILING DATE: 2002-02-27  
 PRIOR APPLICATION NUMBER: 60/364,000  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/404,821  
 PRIOR FILING DATE: 2002-08-08  
 PRIOR APPLICATION NUMBER: 60/334,526  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,409  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/364,227  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/334,027  
 PRIOR FILING DATE: 2001-11-28  
 PRIOR APPLICATION NUMBER: 60/331,641  
 PRIOR FILING DATE: 2001-11-20  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 673  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO: 238  
 LENGTH: 1095  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-287-226-238

Query Match 41.5%; Score 52; DB 15; Length 1095;  
 Best Local Similarity 40.0%; Pred. No. 24;  
 Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIIAGRRLRMRMGDQNGELEASAK 25  
 Db 914 LVSSITSGLLTIGDRGGALDAAK 938

RESULT 11  
 US-10-287-226-246  
 ; Sequence 246, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agee, Michele L.,  
 ; APPLICANT: Alsobrook, John P.,  
 ; APPLICANT: Berghs, Constance,  
 ; APPLICANT: Boldog, Ference,  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.,  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisner, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramcov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malyankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Mille, Isabelle,

; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatiana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastelli, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothenberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,  
 ; APPLICANT: Spyrek, Kimberley A.,  
 ; APPLICANT: Taupier, Jr., Raymond J.,  
 ; APPLICANT: Vernet, Corine A.M.,  
 ; APPLICANT: Zerhusen, Bryan D.,  
 ; APPLICANT: Zhong, Mei  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; CURRENT APPLICATION NUMBER: US/10/287,226  
 ; CURRENT FILING DATE: 2002-11-04  
 ; FILE REFERENCE: 21402-480C  
 ; CURRENT APPLICATION NUMBER: US/10/287,226  
 ; CURRENT FILING DATE: 2002-11-04  
 ; PRIOR APPLICATION NUMBER: 60/334,421  
 ; PRIOR FILING DATE: 2001-11-30  
 ; PRIOR APPLICATION NUMBER: 60/354,392  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/360,148  
 ; PRIOR FILING DATE: 2002-02-27  
 ; PRIOR APPLICATION NUMBER: 60/364,000  
 ; PRIOR FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: 60/404,821  
 ; PRIOR FILING DATE: 2002-08-20  
 ; PRIOR APPLICATION NUMBER: 60/334,526  
 ; PRIOR FILING DATE: 2002-11-30  
 ; PRIOR APPLICATION NUMBER: 60/364,409  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/364,227  
 ; PRIOR FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: 60/334,027  
 ; PRIOR FILING DATE: 2001-11-28  
 ; PRIOR APPLICATION NUMBER: 60/331,641  
 ; PRIOR FILING DATE: 2001-11-20  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 673  
 ; SOFTWARE: CuraseqList version 0.1  
 ; SEQ ID NO: 246  
 ; LENGTH: 1095  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-287-226-246

Query Match 41.6%; Score 52; DB 15; Length 1095;  
 Best Local Similarity 40.0%; Pred. No. 24;  
 Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAILAGRRLRMRMGDQNGELEASAK 25  
 Db 914 LVSSITSGLLTIGDRGGALDAAK 938

RESULT 12  
 US-10-287-226-668  
 ; Sequence 668, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agee, Michele L.,  
 ; APPLICANT: Alsobrook, John P.,  
 ; APPLICANT: Berghs, Constance,  
 ; APPLICANT: Boldog, Ference,  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.,  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisner, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramcov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malyankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Mille, Isabelle,

APPLICANT: Gorman, Linda,  
Gerlach, Valerie,  
APPLICANT: Ji, Weizhen,  
APPLICANT: Kekuda, Ramesh,  
APPLICANT: Khrantsov, Nikolai,  
APPLICANT: Li, Li,  
APPLICANT: Malyankar, Uriel M.,  
APPLICANT: MacDougall, John R.,  
APPLICANT: Mezes, Peter S.,  
APPLICANT: Miller, Charles E.,  
APPLICANT: Millet, Isabelle,  
APPLICANT: Ooi, Chean Eng,  
APPLICANT: Ort, Tatiana,  
APPLICANT: Padigaru, Muralidhara,  
APPLICANT: Pattarajan, Meera,  
APPLICANT: Rastelli, Luca,  
APPLICANT: Rieger, Daniel K.,  
APPLICANT: Rothenberg, Mark E.,  
APPLICANT: Shenoy, Suresh G.,  
APPLICANT: Spaderna, Steven K.,  
APPLICANT: Spytek, Kimberley A.,  
APPLICANT: Taupier, Jr., Raymond J.,  
APPLICANT: Verner, Corine A.M.,  
APPLICANT: Zerhusen, Bryan D.,  
APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-480C  
CURRENT APPLICATION NUMBER: US10/287,226  
CURRENT FILING DATE: 2002-11-04  
PRIOR APPLICATION NUMBER: 60/334,421  
PRIOR FILING DATE: 2001-11-30  
PRIOR APPLICATION NUMBER: 60/354,392  
PRIOR FILING DATE: 2002-02-04  
PRIOR APPLICATION NUMBER: 60/360,148  
PRIOR FILING DATE: 2002-02-17  
PRIOR APPLICATION NUMBER: 60/364,000  
PRIOR FILING DATE: 2002-03-13  
PRIOR APPLICATION NUMBER: 60/404,821  
PRIOR FILING DATE: 2002-08-20  
PRIOR APPLICATION NUMBER: 60/334,526  
PRIOR FILING DATE: 2001-11-30  
PRIOR APPLICATION NUMBER: 60/354,409  
PRIOR FILING DATE: 2002-02-04  
PRIOR APPLICATION NUMBER: 60/364,227  
PRIOR FILING DATE: 2002-03-13  
PRIOR APPLICATION NUMBER: 60/334,027  
PRIOR FILING DATE: 2001-11-28  
PRIOR APPLICATION NUMBER: 60/331,641  
PRIOR FILING DATE: 2001-11-20  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 673  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO: 658  
LENGTH: 1100  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-10-287-226-668

Query Match 41.6%; Score 52; DB 15; Length 1100;  
Best Local Similarity 40.0%; Pred. No. 24;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIIAGRIRMLGDQFNGELEASAK 25  
Db 919 LVSSLTSGLLTIGDRFGALDAAK 943

RESULT 13  
US-10-307-817-84  
Sequence 84, Application US10307817  
Publication No. US20040058338A1  
GENERAL INFORMATION:  
APPLICANT: Agee et al.

APPLICANT: Agee et al.  
FILE REFERENCE: 21402-502C  
CURRENT APPLICATION NUMBER: US10/307,817  
CURRENT FILING DATE: 2002-12-02  
NUMBER OF SEQ ID NOS: 682  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO: 92  
LENGTH: 1101  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-307-817-84

Query Match 41.6%; Score 52; DB 15; Length 1101;  
Best Local Similarity 40.0%; Pred. No. 24;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIIAGRIRMLGDQFNGELEASAK 25  
Db 920 LVSSLTSGLLTIGDRFGALDAAK 944

RESULT 14  
US-10-307-817-92  
Sequence 92, Application US10307817  
Publication No. US20040058338A1  
GENERAL INFORMATION:  
APPLICANT: Agee et al.  
FILE REFERENCE: 21402-502C  
CURRENT APPLICATION NUMBER: US10/307,817  
CURRENT FILING DATE: 2002-12-02  
NUMBER OF SEQ ID NOS: 682  
SOFTWARE: CuraseqList version 0.1  
SEQ ID NO: 92  
LENGTH: 1101  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-307-817-92

Query Match 41.6%; Score 52; DB 15; Length 1101;  
Best Local Similarity 40.0%; Pred. No. 24;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIIAGRIRMLGDQFNGELEASAK 25  
Db 920 LVSSLTSGLLTIGDRFGALDAAK 944

RESULT 15  
US-10-287-226-226  
Sequence 226, Application US10287226  
Publication No. US20040086875A1  
GENERAL INFORMATION:  
APPLICANT: Agee, Michele L.,  
APPLICANT: Alsobrook, John P.,  
APPLICANT: Berghs, Constance,  
APPLICANT: Boldog, Ference,  
APPLICANT: Burgess, Catherine E.,  
APPLICANT: Chant, John S.,  
APPLICANT: Chaudhuri, Amitabha,  
APPLICANT: DiPippo, Vincent A.,  
APPLICANT: Edinger, Shlomit R.,  
APPLICANT: Eisen, Andrew,  
APPLICANT: Ellerman, Karen,  
APPLICANT: Gangoli, Esha A.,  
APPLICANT: Gorman, Linda,  
APPLICANT: Gerlach, Valerie,  
APPLICANT: Ji, Weizhen,  
APPLICANT: Kekuda, Ramesh,  
APPLICANT: Khrantsov, Nikolai,  
APPLICANT: Li, Li,  
APPLICANT: Malyankar, Uriel M.,  
APPLICANT: MacDougall, John R.,  
APPLICANT: Agee et al.

APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Pattarajan, Meera,  
 APPLICANT: Rascelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark E.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spadera, Steven K.,  
 APPLICANT: Spyrek, Kimberley A.,  
 APPLICANT: Taupier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zethusen, Bryan D.,  
 APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-480C  
 CURRENT APPLICATION NUMBER: US10/287,226  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,392  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/350,148  
 PRIOR FILING DATE: 2002-02-27  
 PRIOR APPLICATION NUMBER: 60/364,000  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/404,821  
 PRIOR FILING DATE: 2002-08-20  
 PRIOR APPLICATION NUMBER: 60/334,526  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,409  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/354,227  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/334,027  
 PRIOR FILING DATE: 2001-11-28  
 PRIOR APPLICATION NUMBER: 60/331,641  
 PRIOR FILING DATE: 2001-11-20  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO: 226  
 LENGTH: 1.101  
 TYPE: PRF  
 ORGANISM: Homo sapiens  
 US-10-287-226-226

Query Match 41 6%; Score 52; DB 15; Length 1101;  
 Best Local Similarity 40.0%; Pred. No. 24;  
 Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 IVAIAITGRMLGOFNGBLEASAK 25  
 :|: :|: :|: :|: :|:  
 Db 920 LVSSLTSGLTIGDRFGALDAAK 944

Search completed: November 11, 2004, 02:43:00  
 Job time: 30.1217 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

Copyright

OM protein - Protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 6.37736 Seconds  
[without alignments] 392.268 Million cell updates/sec

Title: US-10-092-750-44

Perfect score: 125

Sequence: 1 IVAIIAGRLMLQFNGELEASARN 26

Scoring table: BLOSUM62

GapP 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.6	1100	2 A35007	ATP citrate (pro-S)-lyase
2	52	41.6	1105	2 S21173	ATP citrate (pro-S)-lyase - human
3	52	41.6	3519	2 S43048	ATP citrate synthase
4	49	39.2	579	2 T02574	hypothetical prote
5	47.5	38.0	146	2 H75111	hypothetical prote
6	47	37.6	140	2 C86800	prophage pi3 prote
7	47	37.6	140	2 B86683	prophage pi1 prote
8	47	37.6	657	2 JC7767	isocamylase (EC 3.2)
9	46	36.8	360	2 AC2436	polyamine-binding
10	45.5	36.4	148	1 H71021	hypothetical prote
11	45	36.0	132	2 F70557	hypothetical prote
12	45	36.0	297	2 AB0431	lysR family trans-
13	45	36.0	996	2 F86410	protein F3M18-12 [
14	45	36.0	2703	2 H81193	hemagglutinin/hemo
15	44.5	35.6	44	2 B69677	phosphatase (Bape)
16	44.5	35.6	144	4 I51936	hypothetical EC0/A
17	44.5	35.6	308	2 D33858	thiamin biosynthes
18	44	35.2	265	2 EB3599	hypothetical prote
19	44	35.2	361	2 AD3198	NN8-4AG - human (f
20	44	35.2	412	2 G02453	series-type D-Ala-
21	44	35.2	479	2 A64117	conserved hypothet
22	44	35.2	593	2 AB3089	hypothetical prote
23	44	35.2	593	2 E38197	hypothetical prote
24	44	35.2	1099	2 T18713	hypothetical prote
25	44	35.2	1106	2 T29496	probable nuclear P
26	44	35.2	1159	2 S6562	protein-tyrosine k
27	43.5	34.8	112	2 B24773	protein-tyrosine k
28	43.5	34.8	113	2 C24773	protein-tyrosine k
29	43.5	34.8	137	2 D24773	protein-tyrosine k

## ALIGNMENTS

RESULT 1

A35007 ATP citrate (EC 4.1.3.8) - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Jul-1990 #text\_change 16-Aug-2004

C;Accession: A35007 ; A35700 ; R;Elshourbagy, N.A.; Near, J.C.; Kmetz, P.J.; Sathé, G.M.; Southan, C.; Strickler, J.E.; J. Biol. Chem. 265, 1430-1435, 1990

A;Title: Rat ATP citrate-lyase. Molecular cloning and sequence analysis of a full-length A;Reference: A35007 ; MUID:90101095 ; PMID:2295639

A;Accession: A35700

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1100

A;Cross-references: UNIPROT:PI6638; GB:J05210; NID:9949989; PIDN:AAA74463.1; PMID:g203490

R;Ramakrishna, S.; D'Angelo, G.; Benjamin, W.B. Biochemistry 29, 7617-7624, 1990

A;Title: Sequence of sites on ATP-citrate lyase and phosphatase inhibitor 2 phosphorylation sites

A;Reference number: A35700 ; MUID:91104719 ; PMID:2176822

A;Accession: A35700

A;Status: preliminary

A;Molecule type: protein

A;Residues: 418-59 <RAM>

C;Superfamily: ATP citrate synthase

C;Keywords: carbon-carbon lyase; coenzyme A; oxo-acid-lyase; phosphoprotein

Query Match Score 52; DB 2; Length 1100;

Best Local Similarity 40.0%; Pred. No. 8.3%; Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAILAIGRLMLQFNGELEASAK 25

Db 919 LVSSLGSLLTGDRGGALDAAK 943

RESULT 2

S21173 ATP citrate (pro-S)-lyase - human

C;Species: Homo sapiens (man)

C;Accession: S21173 ; A;Cross-references: UNIPROT:P53396; EMBL:X64330; NID:928934 ; PMID:CAA5614.1; PMID:g28935

C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004

C;Elshourbagy, N.A.; Near, J.C.; Kmetz, P.J.; Wells, T.N.C.; Groot, P.H.E.; Saxty, B.A.; Eur. J. Biochem. 204, 491-499, 1992

A;Title: Cloning and expression of a human ATP-citrate lyase cDNA.

A;Reference number: S21173 ; MUID:92174902 ; PMID:1371749

A;Accession: S21173

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1105 <ELS>

A;Cross-references: UNIPROT:PI6638; GB:J05210; NID:9949989; PIDN:AAA74463.1; PMID:g203490

C;Superfamily: ATP-citrate synthase

Query Match Score 41.6%; DB 2; Length 1105;  
Best Local Similarity 40.0%; Pred. No. 8.3.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VAIIAQRRLMIGDQFNGELLEASAK 25  
Db 924 LVSSLTSGLTLIGDREFGALDAAK 948

RESULT 3  
S43048 polyketide synthase type I - Streptomyces antibioticus  
N;Contains: acyl carrier protein; acyltransferase; ketoacylsynthase; ketoreductase; thioesterase; thiocarbonyl transferase; thioredoxin reductase  
C;Species: Streptomyces antibioticus  
C;Date: 13-Jan-1995 #sequence\_revision 06-Dec-1996 #text\_change 16-Aug-2004  
R;Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.  
A;Reference number: S43048  
A;Accession: S43048  
A;Molecule type: DNA  
A;Residues: 1-3519 <SWA>  
A;Cross-references: UNIPROT:Q07017; EMBL:LO9654; NID:S153407; PID:AAA19695.1; PID:gi1534  
R;Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.  
Mol. Gen. 242, 358-362, 1994  
A;Title: Characterisation of a Streptomyces antibioticus gene encoding a type I polyketide synthase  
A;Reference number: S41729; PMID:9150470; MUID:9150470  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-683-3238, 3273-3303, 'K', 3305-3407, 'T', 3409-3462, 'Y', 3464-3516, 'E', 3518-3519  
A;Cross-references: EMBL:LO9654  
C;Genetics:  
A;Start codon: GTG  
C;Keywords: antibiotic biosynthesis; carrier protein  
F:570-462/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
F:570-851/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>  
F:1202-1381/Domain: short-chain alcohol dehydrogenase homology <SDA1>  
F:1489-1560/Domain: acyl carrier protein homology <ACI1>  
F:1708-2111/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>  
F:2221-2502/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>  
F:3143-3214/Domain: short-chain alcohol dehydrogenase homology <AC22>  
F:3305-3500/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>  
Query Match Score 41.6%; DB 2; Length 3519;  
Best Local Similarity 40.0%; Pred. No. 30;  
Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 2 VAIAGRLRMLGDQFNGELLEASAKN 26  
Db 701 VSLPAGRVRVITMLEEFGRSLVAVN 725

RESULT 4  
T02574 hypothetical protein At2g39280 [imported] - Arabidopsis thaliana  
N;Alternate names: hypothetical protein T16524.8  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: T02574; D84815  
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, A;  
A;Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.  
A;Reference number: Z14679  
A;Accession: T02574  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-579 <ROU>  
A;Cross-references: UNIPROT:080953; EMBL:AC004697; NID:g3402671; PID:g3402693  
A;Experimental source: cultiver Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Umayan, L.; Tailor, L.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayan, L.; Tailor, L.

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A844242; MUID:20083487; PMID:10617197  
A;Accession: D8815  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-579 <STO>  
A;Cross-references: GB:AE002092; NID:g3402693; PIDN: AAC28996.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g39280; T16B24.8  
A;Map position: 2  
A;Introns: 219/3; 279/3; 317/3; 360/2; 382/3; 400/1; 428/1  
Query Match Score 39.2%; DB 2; Length 579;  
Best Local Similarity 43.3%; Pred. No. 12;  
Matches 10; Conservative 8; Indels 0; Gaps 0;  
5; Mismatches 8;  
Qy 4 IAAGRNLMGDQFNGELLEASAKN 26  
Db 118 LMSVVRVMKGDSNTGEQAKLN 140

RESULT 5  
H7511 hypothetical protein PAB1905 - Pyrococcus abyssi (strain Orsay)  
A;Species: Pyrococcus abyssi  
C;Accession: H7511  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;anonymous. Genoscope submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A;Reference number: A75001  
A;Accession: H7511  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-146 <XAW>  
A;Cross-references: UNIPROT:Q9V0U4; GB:AJ248285; PID:gi5458067; PID:CAL096836; PID:CA49605;  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB1905  
C;superfamily: Escherichia coli hypothetical protein b3356  
Query Match Score 38.0%; DB 2; Length 146;  
Best Local Similarity 45.3%; Pred. No. 4.7;  
Matches 10; Conservative 6; Indels 1; Gaps 1;  
Qy 1 IAIAGRLRMLGDQFNGELLEASAKN 21  
Db 3 IMTIVRGRVTRVTEEQFIGRIE 24

RESULT 6  
C86800 prophage p13 protein 30 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C;Species: Lactococcus lactis subsp. lactis  
C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: C86800  
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaijon, O.; Malarme, K.; Weissenbach, J.; Ehrlic Genome Res. 11, 731-753, 2001  
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. C;Accession: C86800  
A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: C86800  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-140 <STO>  
A;Cross-references: UNIPROT:QCCFR2; GB:AE005176; PID:gi12724390; PID:AAK05501.1; GSPDB:GR  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: pi30

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 10 RMLGDOFNGEASAKN 26  
 Db 7 RIIGDYZVNGRLEARIKS 23

C;Accession: AC2463  
 R;Kaneo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal

A;Reference number: AB1807; MUID:21595225; PMID:11759840

A;Accession: AC2436  
 A;Status: Preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-360 <STOP>  
 A;Cross-references: UNIPROT:QYMM93; GB:BA000019; PIDN:BAB76742.1; PIDN:917134181; GSPDB:Gr

A;Experimental source: strain PCC 7120  
 A;Genetics:  
 C;Superfamily: Escherichia coli spermidine/putrescine-binding protein

Query Match 36.8%; Score 46; DB 2; Length 360;  
 Best Local Similarity 45.5%; Pred. No. 22;  
 Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 4 IIAGRMLGDQFNGEASAK 25  
 Db 185 VIGAVIQLMLGYSNSNEAQIK 206

RESULT 10  
 H71021  
 hypothetical protein PH1469 - Pyrococcus horikoshii  
 Pyrococcus horikoshii  
 C;Species: Pyrococcus horikoshii  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 12-Jul-2004  
 C;Accession: H71021  
 R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.; Res. 5, 55-76, 1998  
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-148 <KRW>  
 A;Cross-references: UNIPROT:OS9138; GB:AP000006; PIDN:BAA30576.1; PID:9325;

A;Experimental source: strain OT3  
 A;Note: This accession replaces an interim accession for a sequence replaced by GenBank  
 C;Genetics:

Query Match 36.4%; Score 45.5; DB 1; Length 148;  
 Best Local Similarity 40.9%; Pred. No. 9.8;  
 Matches 9; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

Db 1 IVIAAGRLRML-GDQFNGE 21  
 5 IVTVVKGKVRWVEQPIGRIE 26

RESULT 11  
 F70557  
 hypothetical protein Rv1616 - Mycobacterium tuberculosis (strain H37RV)  
 Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Accession: F70557  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Ratnarekam, M.A.; Rogers, J.; Rutte, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295937; PMID:98344230

A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-132 <COL>  
 A;Cross-references: UNIPROT:O06133; GB:295554; PIDN:CA08893.1

RESULT 9  
 AC2416  
 polyamine-binding protein of polyamine ABC transporter all5043 [imported] - Nostoc sp.  
 C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

:Experimental source: strain H37RV  
 :Genetics: Rv1616

Query Match Score 45; DB 2; Length 132;  
 Best Local Similarity 58.8%; Pred. No. 10;  
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 RESULT 14  
 H81193

Y 7 GRLMLGQFNGELAES 23  
 b 55 GGLMIDLIHGEAAS 71

7 GRLMLGQFNGELAESAKN 26  
 Gene: Rv1616  
 C:Species: Neisseria meningitidis [imported] - Neisseria meningitidis (strain H37RV)  
 C:Accession: H81193  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C:Tittle: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; PMID:20175755; PMID:1070307  
 A:Accession: H81193  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2703 <TET>  
 A:Cross-references: UNIPROT:Q9KOT0; GB:AE002405; GB:AE002098; PIDN:97225708; PIDN:AAF40927  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: XNB0493

Query Match Score 45; DB 2; Length 2703;  
 Best Local Similarity 39.4%; Pred. No. 2.8e+02;  
 Matches 13; Conservative 5; Mismatches 3; Indels 12; Gaps 2;

Y 6 AGRMLGQFN---GELPA-----SAKN 26  
 b 552 AAKLRYSGDSEFNNTYTGKLQAHDLAVNTQAKN 584

7 GRLMLGQFNGELAESAKN 26  
 Gene: YPO3545  
 C:Species: Bacillus subtilis  
 C:Accession: B96677  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Tittle: phosphatase (Rape) regulator phrE - Bacillus subtilis  
 A:Authors: Poulier, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Galizzi, A.; Galleni, C.; Kunz, F.; Ogasawara, N.; Noszer, I.; Albertini, A.M.; Alioni, G.; Azevedo, V.; Bertorelli, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Callewell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Koetter, P.; Konigstein, G.; Krogh, S.; Kurata, K.; Lapic, S.; Lardinois, M.; Levine, A.; Liu, H.; Masuda, S.; Mauel, J.; Mauel, S.; Portetelle, Y. M.; Ogawa, K.; Ogiwara, A.; Onaga, B.; Park, S.H.; Patro, V.; Pohl, T.M.; Portetelle, Y.; Rieger, M.; Rivolta, C.; Rocha, E.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, Akeuchi, M.; Tamaki, A.; Tanaka, P.; Terpstra, P.; Togomi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshihara, H.; Danchin, A.; Danchin, A.; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; PMID:98044033; PMID:9384377  
 A:Accession: B96677  
 A:Molecule type: DNA  
 A:Residues: 1-44 <KUN>  
 A:Cross-references: UNIPROT:Q32025; GB:Z99117; PIDN:G2634966; PIDN:CA14525.1  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: phrE

Query Match Score 45; DB 2; Length 44;  
 Best Local Similarity 33.3%; Pred. No. 3.8;  
 Matches 8; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Y 3 ATTAQRMLGQFNGELAESAKN 26  
 Gene: F3M18.12  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Accession: F86410  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Tittle: protein kinase P3M18.12 [imported] - Arabidopsis thaliana  
 A:Authors: Boller, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Hughes, N.F.; Huizar, L.; Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.; Kondo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, J.R.; Shian, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:1110712  
 A:Accession: F86410  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-996 <STO>  
 A:Cross-references: UNIPROT:Q9SGP2; GB:AEE005172; PIDN:96560764; PIDN:AAF16764.1; GSDB:GN  
 A:Experimental source: strain not shown  
 C:Genetics:  
 A:Gene: F3M18.12

Query Match Score 45; DB 2; Length 996;  
 Best Local Similarity 45.0%; Pred. No. 95;  
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 11 AVLIG-LAFFGSMYNGEMBA SRN 33

Search completed: November 10, 2004, 14:52:14  
Job time : 7.37736 sec<sub>E</sub>

This Page Blank (uspto)

Copyright	Gencore version 5.1.6 (c) 1993 - 2004 Compugen Ltd.	OM protein - protein search, using sw model	Run on: November 10, 2004, 13:38:57 ; Search time 34.614 Seconds 'without alignments'; 431.938 Million cel. updates/sec	Scoring table: BLOSUM62 Gappext 10.0 , Gappext 0.5	Searched: 1825181 seqs, 575374646 residues	Total number of hits satisfying chosen parameters:	1825181.	ALIGMENTS
Minimum DB seq length:	0	Maximum DB seq length:	2000000000	Post-processing: Minimum Match 0%	Maximum Match 100%	Listing first 45 summaries		
Database :	Uniprot_02: 1: uniprot_sprot: 2: uniprot_trembl: *					Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
						SUMMARIES		
Result No.	Score	Query	Match	Length	DB ID	Description		
1	108	Q8K186	157	2	Q8EAN7	Q8k186 mus musculus		
2	53	Q8EAN7	393	2	Q99K05	Q8ean7 shewanella		
3	52	Q99K05	195	2	Q8VDM8	Q8905 mus musculus		
4	52	Q8VDM8	592	2	Q7VZ19	Q8vdm8 mus musculus		
5	52	Q7VZ19	411	6	Q7WDA8	Q7vz19 bordetellla		
6	52	Q7WDA8	682	2	Q8N9C4	Q7wda8 bordetellla		
7	52	Q8N9C4	411	6	Q8V1Q1	Q8n9c4 homo sapien		
8	52	Q8V1Q1	701	2	Q91922	Q8v1q1 rattus norv		
9	52	Q91922	1091	1	Q165338	Q91922 mus musculus		
10	52	Q165338	1100	1	Q53496	P165338 rattus norv		
11	52	Q53496	1101	1	Q5L6_HUMAN	P53496 homo sapien		
12	52	Q5L6_HUMAN	411	6	Q6DG67	Q5l6_human		
13	51	Q6DG67	3519	1	Q6D667	O6dg67 streptomyce		
14	49	Q6D667	1092	2	Q700Q0	O6d667 brachydanio		
15	49	Q700Q0	248	2	CAF32984	Q700q0 pseudomonas		
16	49	CAF32984	278	2	Q7PMH0	Caf32984 pseudomonas		
17	49	Q7PMH0	314	2	Q7PTU4	Q7pmh0 anophelis		
18	49	Q7PTU4	391	2	Q7JNC6	Q7pmh0 anophelis		
19	49	Q7JNC6	391	2	Q7Q0U7	Q7jnc6 drosophilal		
20	49	Q7Q0U7	579	2	Q80953	Q7q0u7 arabidopsis		
21	49	Q80953	607	1	UYRC_PSEPK	Q80953 pseudomonas		
22	49	UYRC_PSEPK	1086	2	Q7KRA9	Q80953 pseudomonas		
23	49	Q7KRA9	1086	2	AAM0940	Q80953 pseudomonas		
24	49	AAM0940	1112	2	Q7KNB5	Aam0940 drosophilal		
25	49	Q7KNB5	1112	2	AD34754	Q7knb5 drosophilal		
26	49	AD34754	391	2	Q7Q0U7	Q7knb5 anophelis		
27	49	Q7Q0U7	1118	2	Q9KIV3	Q7q0u7 anophelis		
28	48	Q9KIV3	3816	2	Q9KIV4	Q9kiv3 streptomyce		
29	48	Q9KIV4	4150	2	Q93RK2	Q9kiv4 streptomyce		
30	48	Q93RK2	424	2	Q9AD52	Q93rk2 haemophilus		
31	48	Q9AD52	723	2	Q9BIX2	Q9ad52 streptomyce		
		Q9BIX2	1054	2	Q6PGM5	Q9bix2 gallus gallus		
		Q6PGM5	1223	2		Q6pgm5 gloeobacter		

Aah56933 mus musculu  
Q8chh7 mus musculu  
Pyrococcus  
Q89b89 bradyrhizob  
O77ij0 lactococcus  
Q9ayx5 lactococcus  
Q9az68 lactococcus  
Q9azm9 bacteriopha  
Q9azt6 bacteriopha  
Q9cfz2 lactococcus  
Q9cfz2 lactococcus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
NCBI\_TaxID=10090;  
RN

SEQUENCE FROM N.A.  
TISSUE= Mammary gland;  
MEDLINE=238825 ; PubMed=12477932;

RA Strassberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Bhat N.K.,  
RA Altshul S.F., Zeeberg B., Buerk K.H., Schaefter C.P., Bhat N.K.,  
RA Diatchenko L., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Hopkins R.P., Maruska K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Potschuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquai-Laino N.A., Abramson R.D., Mulahay S.J.,  
RA Bosak S.A., McEvani P.J., McKernan K.J., Malek J.A., Guarante P.H.,  
RA Richards D., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.R., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.,  
RA Rodriguez A.C., Grimwade J., Schmutz J., Myers R.M., Scheiner J.B.,  
RA Krzywinski M.I., Skalska U., Smailius D.E., Scheiner J.B., Jones S.J., Marrs M.A.,  
RA SEQUENCE FROM N.A.  
TISSUE= Mammary gland;

Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL: BC027668; AAH27668.1; - .  
NON-TER 1 .  
SQ SEQUENCE 157 AA; 17078 MW; 92FF95902186B55347 CRC64;  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN

SEQUENCE FROM N.A.  
TISSUE= Mammary gland;

Query Match 86.4%; Score 108; DB 2; Length 157;  
Best Local Similarity 88.0%; Pred. No. 2e-08;  
Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VALIAGDRMUDQFQNGPLEAKN 26  
Db 43 VALIVGRLRILGDQFQNGPLEAKN 67

**RESULT 2**

Q8EAN7 PRELIMINARY; PRT; 393 AA.

ID Q8EAN7 ID ; PRELIMINARY; PRT; 393 AA.

AC Q8EAN7 ID ; PRELIMINARY; PRT; 393 AA.

DT 01-MAR-2003 (TREMBrel. 23, Created)

DT 01-MAR-2003 (TREMBrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBrel. 26, Last annotation update)

DE Iron-sulfur cluster-binding protein.

OrderedLocusNames=SG3861;

GN Shewanella oneidensis.

OS Shewanella oneidensis.

OC Bacteria; Proteobacteria; Alteromonadales;

OC Alteromonadaceae; Shewanellales.

[1]

RN PubMed=22297686; PubMed=12368813; DOI=10.1038/nbt749;

RN MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;

RN Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methé B.A., Heidelberg J.F., Paulsen K.E., Nelson W.C., Clayton R.A., Meyer T., Tsaprin A., Scott J., Beanan M.J., Brinkac M., Kolonay J.F., Madupu R., Dodson R.J., Durkin A.S., Haft D.H., Kolenay J.F., Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M., Vamathevan J.J., Weidman J.P., Impraim M., Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.J., Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O., Venter J.C., Nelson K.H., Fraser C.M., Smith J.V., Genomic sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis." Nat. Biotechnol. 20:1118-1123 (2002).

RN EMBL; AB05819; ANNS6837.1; -.

DR TIGR; SG3861; -.

DR GO:0005489; F:electron transporter activity; IEA.

DR GO:0005506; F:iron ion binding; IEA.

DR GO:0006118; F:electron transport; IEA.

DR InterPro; IPR001450; 4Fe4S ferredoxin.

DR InterPro; IPR004453; FeSClus binding.

DR InterPro; IPR009051; Helical\_ferrredoxin.

DR Pfam; PF00037; Ferl\_1.

DR TIGRFAMs; TIGRF00276; FeSClus binding; 1.

DR PROSITE; PS0198; 4Fe4S\_FERRDOXIN\_1.

KW SEQUENCE; 4Fe-4S; Complete proteome; Iron; Iron-sulfur; Metal-binding.

SQ 393 AA; 43580 MW; 65PE082721B338 CRC64;

**RESULT 3**

Q99K05 PRELIMINARY; PRT; 195 AA.

ID Q99K05 ID ; PRELIMINARY; PRT; 195 AA.

AC Q99K05 ID ; PRELIMINARY; PRT; 195 AA.

DT 01-JUN-2001 (TREMBrel. 17, Created)

DT 01-OCT-2003 (TREMBrel. 17, Last sequence update)

DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)

DB ACly protein (Fragment).

GN Mus musculus (Mouse).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.

CX NCBI\_TaxID=10030;

RN LIRARLKLGDDQINSBLVA 141

RN SEQUENCE FROM N.A.

RC STRAIN=mix FVB/N.

RC TISSUE=Mammary tumor.

RC MEDLINE=22388257; PubMed=12477932;

RC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Sheinen C.M., Schuler G.D., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Casavant T.L., Grouse L.H., Derge J.G., Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Brownstein M.J., Usdin T.B., Zeeberg B., Buetow K.H., Schaefer C.F., Prange C., Carninci P., Gunaratne P.H., Raha S.S., Loqueland N.A., Abramson R.D., Mullanay S.J., Bosak S.A., McKernan K.J., Malek J.A., Hale S., Sodergren E.J., Lu X., Gibbs R.A., Richards S., Worley K.C., Hale S., Sodergren E.J., Lu X., Gibbs R.A., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schein J.E., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN SEQUENCE FROM N.A.

RC STRAIN=mix FVB/N.

RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old; Strausberg R.; Submitted (NAR-2001) to the EMBL/GenBank/DBJ databases.

RN Submitted (NAR-2001) to the EMBL/GenBank/DBJ databases.

RN EMBL; BC00533; AAH5533.1; -.

DR MGD; MGI:103251; Acly.

DR GO: GO:004108; F:citrate (Si)-synthase activity; IEA.

DR GO: GO:000099; P:tricarboxylic acid cycle; IEA.

DR InterPro; IPR002020; Citrate\_synth.

FT NON\_TER 1

SEQUENCE 195 AA; 21693 MW; 68127314F9BB3404 CRC64;

Query Match Score 52; DB 2; Length 195;

Best Local Similarity 41.6%; Score 52; DB 2; Length 195;

Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAILAGRLRMLGDFNGELEASAK 25

DB 14 LVSSLSGLITGREGALDAAK 38.

**RESULT 4**

Q8VDM8 PRELIMINARY; PRT; 592 AA.

ID Q8VDM8 ID ; PRELIMINARY; PRT; 592 AA.

AC Q8VDM8 ID ; PRELIMINARY; PRT; 592 AA.

DT 01-MAR-2002 (TREMBrel. 20, Created)

DT 01-MAR-2002 (TREMBrel. 20, Last sequence update)

DT 01-MAR-2004 (TREMBrel. 26, Last annotation update)

DB Acly protein (Fragment).

GN Name=Acly;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.

CX NCBI\_TaxID=10030;

RN SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;

RC MEDLINE=22388257; PubMed=12477932;

RC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Sheinen C.M., Schuler G.D., Schaefer C.F., Prange C., Carninci P., Gunaratne P.H., Raha S.S., Loqueland N.A., Abramson R.D., Mullanay S.J., Bosak S.A., McKernan K.J., Malek J.A., Hale S., Sodergren E.J., Lu X., Gibbs R.A., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	KW	Complete proteome.
RT		SQ	SEQUENCE FROM N.A.
RL	SEQUENCE FROM N.A.	Query Match	41.6%; Score 52; DB 2; Length 692;
RN	STRAIN=FVB/N; TISSUE=Mammary tumor. C3; STRAIGHT=1; Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.	Best Local Similarity	48.0%; Pred. No. 40;
RP		Matches	6; Mismatches 5; Indels 2; Gaps 1;
RC		Qy	1 IVAILAGRLRMLGDQFN-GELEAS 23
RA		Db	180 LVSLATGGLRVLGDSFWWGQLGAS 204
RL	EMBL: BC021502; AAH21502.1; -.	RESULT 6	
DR	MGD; MGI:103251; Acyl.	OTWDAB	PRELIMINARY;
DR	GO; GO:0003824; F:catalytic activity; IEA.	ID	PRT; 693 AA.
DR	GO; GO:0004108; F:citrate (S)-synthase activity; IEA.	OTWDAB;	
DR	GO; GO:0008152; P:metabolism; IEA.	AC	
DR	GO; GO:0006099; P:tricarboxylic acid cycle; IEA.	OTWDAB;	
DR	InterPro; IPR002020; Citrate synth.	ID	
DR	InterPro; IPR003781; CoA binding.	OTWDAB;	Created]
DR	InterPro; IPR005111; CoA ligase.	DT	01-OCT-2003 (TREMBLrel. 25; Last sequence update)
DR	InterPro; IPR005810; CoA ligase alpha.	DT	01-MAR-2004 (TREMBLrel. 26; Last annotation update)
DR	Pfam; PF02629; CoA binding; 1.	DE	Putative membrane protein.
DR	Pfam; PF00549; Ligase_CoA; 1.	GN	OrderedLocusName=BB5670;
DR	PRINTS; PR01798; SCOCYNSPHASE.	OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus)
DR	PROSITE; PS01216; SUCCINYL_COA_LIG; 1.	OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
DR	PROSITE; PS00399; SUCCINYL_COA_LIG; 2; 1.	OC	Alcaligenaceae; Bordetellae; Bordetellales;
FT	NON_TER 1_592 AA; 65023 MW; 00B455BC52243C92B CRC64; SEQUENCE	RN	NCBITaxonID=518; [1]
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RP	SEQUENCE FROM N.A.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RC	STRAIN=RB50 / ATCC BAA-588;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RX	MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RA	Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R., Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Tarraga A.-M., Temple L., James K.D., Quail M.A., Achtman M., Atkin R., Baker S., Basham D., Bason N., Chevrevach I., Chillingworth T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C., Rabbinkowitsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J., RT
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RT	"Comparative analysis of the genome sequences of <i>Bordetella pertussis</i> , <i>Bordetella parapertussis</i> and <i>Bordetella bronchiseptica</i> "; Nat. Genet. 35:32-40 (2003).
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	PT	PT
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RL	PT
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	EMBL	EMBL: BX640448; CAE3543; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	InterPro; IPR004814; Oligopept_transpt.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	InterPro; IPR004813; Tetrapept_transpt.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	PFAM	PF031169; OPT; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	TIGRFAMS; TIGR00733; OPT_fam; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	TIGRFAMS; TIGR00728; OPT_fam; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	KW	Complete proteome;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	SQ	SEQUENCE 693 AA; 71F130E7BD611A88 CRC64;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RP	SEQUENCE FROM N.A.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RC	STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RX	MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RA	Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R., Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Tarraga A.-M., Temple L., James K.D., Quail M.A., Achtman M., Atkin R., Baker S., Basham D., Bason N., Chevrevach I., Chillingworth T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C., Rabbinkowitsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J., RT
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RT	"Comparative analysis of the genome sequences of <i>Bordetella pertussis</i> , <i>Bordetella parapertussis</i> and <i>Bordetella bronchiseptica</i> "; Nat. Genet. 35:32-40 (2003).
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	EMBL	EMBL: BX640448; CAE3543; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	InterPro; IPR004814; Oligopept_transpt.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	InterPro; IPR004813; Tetrapept_transpt.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	PFAM	PF031169; OPT; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	TIGRFAMS; TIGR00733; OPT_fam; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DR	TIGRFAMS; TIGR00728; OPT_fam; 1.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RN	[1]
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RESULT 7	
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	QBN9C4	PRELIMINARY;
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	ID	PRT; 701 AA.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	AC	
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	OTWDAB;	Created]
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DT	01-OCT-2002 (TREMBLrel. 22; Last sequence update)
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DT	01-MAR-2004 (TREMBLrel. 25; Last annotation update)
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	DE	Hypothetical protein FLJ37765.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	OC	
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	OC	
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	NCBI_TaxID=9606;	
FT	592 AA; 65023 MW; 00B455BC52243C92B CRC64;	RN	[1]

RP	SEQUENCE FROM N.A.	RN [1]
RC	TISSUE=Hippocampus;	SEQUENCE FROM N.A.
RA	PubMed=14702039;	RC TISSUE=Liver;
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibusawa T., Tanaka T., Ishii S., Murakami K., Yasuda T., Iwayanagi T., Nagamura Y., Nagahori K., Sudo H., Hosocri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Kuroda T., Furuya T., Omura Y., Abe K., Komihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., RA Tana J., Kimata M., Watanabe M., Hirata Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., RA Musasino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Mori Y., Momiyama H., Satoh N., Takada S., Terashina Y., Suzuki O., Nakagawa S., Sanch A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kunita S., Fukuzumi Y., Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohnori Y., RA Kawabata R., Hikiji T., Kobayashi N., Inagaki H., Ikeda Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeita K., Senba T., RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., RA Togashi T., Oyama M., Hara H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., RA Nakai K., Nagase T., Nomura Y., Kikuchi H., Masuho Y., Yamashita R., RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., RT "Complete sequencing and characterization of 21,243 full-length human cDNAs"; RT "Complete sequencing and characterization of 21,243 full-length human cDNAs"; Nat. Genet. 36:40-45(2004); DR EMBL; AK050084; BAC04484.1; -.	
DR	HSSP; P003824; JJKU.	GO; GO:0004108; F: citrate (S1)-synthase activity; IEA.
DR	GO; GO:0004152; F: metabolism; IEA.	GO; GO:0006099; F: tricarboxylic acid cycle; IEA.
DR	InterPro; IPR002020; Citrate synth.	InterPro; IPR003781; CoA binding.
DR	InterPro; IPR003781; CoA binding.	InterPro; IPR005811; CoA ligase.
DR	InterPro; IPR005811; CoA ligase.	InterPro; IPR005809; CoA lig. alpha.
DR	Pfam; PF00229; CoA binding.	Pfam; PF00229; CoA binding.
DR	Pfam; PF00549; Ligase CoA; 1.	Pfam; PF00549; Ligase CoA; 1.
DR	PRINTS; PRO1798; SCOCSYNTHASE.	PRINTS; PRO1798; SCOCSYNTHASE.
DR	PROSITE; PS01216; SUCCINYL COA LIG 1; 1.	PROSITE; PS01216; SUCCINYL COA LIG 1; 1.
DR	PROSITE; PS00399; SUCCINYL COA LIG 2; 1.	PROSITE; PS00399; SUCCINYL COA LIG 2; 1.
SQ	SEQUENCE 701 AA; 76445 MW; 4BEC567E3D57852 CRC64;	SEQUENCE 701 AA; 92479 MW; 0717514686408CA0 CRC64;
Qy	1 IVAILAGRLMLGDQNGEELASAK 544	Query Match 41.6%; Score 52; DB 2; Length 701; Best Local Similarity 40.0%; Pred. No. 42; Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0; Gaps 0;
Db	520 LVSSLTSGLLTIDREGGALDAAK 544	DB 670 LVSSLTSGLLTIDREGGALDAAK 694
RESULT 8		
Q8V1Q1	PRELIMINARY;	ACTY_MOUSE ID ACTY_MOUSE STANDARD; PRT; 1091 AA.
AC	Q8V1Q1;	AC Q91V92; DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT	01-MAR-2002 (T-EMBLrel. 20, Last sequence update)	DT 01-MAR-2002 (T-EMBLrel. 26, Last annotation update)
DT	ATP-Citrate lyase (Fragment).	NCBI_TaxID=10116; Name=ACI; OS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
GN		OSIS; Last update
OS		OSIS; Last update
OC		OCIS; Last update
OC		OCIS; Last update
OX		OXIS; Last update
RESULT 9		
Qy	1 IVAILAGRLMLGDQNGEELASAK 544	Query Match 41.6%; Score 52; DB 2; Length 851; Best Local Similarity 40.0%; Pred. No. 50; Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
Db	670 LVSSLTSGLLTIDREGGALDAAK 694	DB 670 LVSSLTSGLLTIDREGGALDAAK 694

DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE ATP-citrat synthase (EC 2.3.3.8) (ATP-citrate (pro-S)-lyase)  
DE (Citrate cleavage enzyme).  
GN Name=AcL1;  
OS Mus musculus (Mouse);  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10050;  
RN [1] SEQUENCE FROM N.A.  
RC STRAIN=LSS, and ISS;  
RX MEDLINE=21363810; PubMed=11471062;  
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,  
RA Beeson M., Gordon L., Bennett B., Johnson T.B., Sikeia J.M.;  
RT "High-throughput sequence identification of gene coding variants  
within alcohol-related QTLs".  
RL Mamm. Genome 12:657-663 (2001);  
RN [2] SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=2238257; PubMed=11477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.D., Collins F.S., Wagner L., Sheimani C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhattacharjee A.,  
RA Hopkins R.P., Jordao S., Moore T., Max J., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Benaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahay S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergran B.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shewchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinberg J., Schmutz J., Skalska U., Marra M.A.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Murinae; Rattus.  
RA Schnarf A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -|- FUNCTION: ATP citrate-lyase is the primary enzyme responsible for  
the synthesis of cytosolic acetyl-CoA in many tissues. Has a  
central role in de novo lipid synthesis. In nervous tissue it may  
be involved in the biosynthesis of acetylcholine. (By similarity).  
CC -|- CATALYTIC ACTIVITY: ATP + Phosphate + acetyl-CoA + oxaloacetate =  
ATP + citrate + CoA.  
CC -|- SUBUNIT: Homotrimer. (By similarity).  
CC -|- SUBCELLULAR LOCATION: Cyttoplasmic (By similarity).  
CC -|- SIMILARITY: In the N-terminal section; belongs to the  
succinate/malate CoA ligase beta subunit family.  
CC -|- SIMILARITY: In the C-terminal section; belongs to the  
succinate/malate CoA ligase alpha subunit family.  
CC -|- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).  
CC -|-  
DR EMBL; AF312052; AAK56081; 1;  
DR EMBL; AF332051; AAK56080; 1;  
DR EMBL; BC056378; AAH56378; 1;  
DR HSSP; P07459; ISCU.  
DR MGD; MGI:103251; AcLY.  
DR InterPro; IPR002020; Citrate synth.  
DR InterPro; IPR003781; CoA binding.  
DR InterPro; IPR005811; CoA Ligase.  
DR InterPro; IPR005810; CoA Lig alpha.  
DR InterPro; IPR005809; CoA binding; 1.  
DR Pfam; PF02629; CoA\_binding; 1.

DR Pfam; PF00549; Ligase\_CoA; 1.  
DR PROSITE; PS01216; SUCCINYL\_COA\_LIG\_1; 1.  
DR PROSITE; PS00399; SUCCINYL\_COA\_LIG\_2; 1.  
DR PROSITE; PS1117; SUCCINYL\_COA\_LIG\_3; 1.  
KW ATP-binding; Lipid synthesis; Phosphorylation; Transferase.  
FT NP\_BIND 691 711 ATP (By similarity).  
FT NP\_BIND 742 768 ATP (By similarity).  
FT DOMAIN 769 779 CoA-binding (potential).  
FT MOD\_RES 455 455 Phosphoserine (by PKA) (By similarity).  
FT ACT\_SITE 750 750 Tele-phosphohistidine intermediate (By similarity).  
FT METAL 708 708 Magnesium (By similarity).  
SQ SEQUENCE 1091 AA; 119727 MW; 650293D027D97DD CRC64;  
Query Match Score 52; DB 1; Length 1091;  
Best Local Similarity 40.0%; Pred. No. 64;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVAIAGRLRMLGDFNGELEASAK 25  
Db 910 LVSSLTSGLLTGDRFFGALDAK 934

RESULT 10  
ACLY RAT STANDARD PRT; 1100 AA.  
ID ACLY RAT ID ACLY RAT P166389;  
AC 15, Created  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-AUG-1990 (Rel. 15, Last annotation update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S)-lyase)  
GN Name=AcLY;  
OS Rattus norvegicus (Rat).  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
RN [1] RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RN [1] RC TISSUE=Liver;  
RX MEDLINE=90110199; PubMed=2295639;  
RA Elshourbagy N.A., Near J.C., Knetz P.J., Sathe G.M., Southan C.,  
RA Strickler J.B., Gross M., Young J.F., Wells T.N.C., Groot P.H.E.;  
RA RT "Rat ATP citrate lyase. Molecular cloning and sequence analysis of a  
full-length cDNA and mRNA abundance as a function of diet, organ, and  
age.";  
RA RL J. Biol. Chem. 265:1430-1435 (1990).  
CC -|- FUNCTION: ATP citrate-lyase is the primary enzyme responsible for  
the synthesis of cytosolic acetyl-CoA in many tissues. Has a  
central role in de novo lipid synthesis. In nervous tissue it may  
be involved in the biosynthesis of acetylcholine. (By similarity).  
CC -|- PHOSPHORYLATION STATE OF ONE OF THE REGULATORY SITES DEPENDS ON THE  
CC -|- AMOUNT OF ENZYME.  
CC -|- SUBUNIT: Homotrimer.  
CC -|- SUBCELLULAR LOCATION: Cyttoplasmic.  
CC -|- PTM: Phosphorylated on two regulatory serines. In nervous tissue it may  
be involved in the biosynthesis of acetylcholine. Molecular cloning and sequence analysis of a  
full-length cDNA and mRNA abundance as a function of diet, organ, and  
age.";  
CC -|- CATALYTIC ACTIVITY: ADP + phosphate + acetyl-CoA + oxaloacetate =  
ATP + citrate + CoA.  
CC -|- ENZYME REGULATION: Major regulation of ATP citrate-lyase activity  
is probably not by phosphorylation/dephosphorylation but by  
altering the amount of enzyme.  
CC -|- PTM: The N-terminus is blocked.  
CC -|- SIMILARITY: In the N-terminal section; belongs to the  
succinate/malate CoA ligase beta subunit family.  
CC -|- SIMILARITY: In the C-terminal section; belongs to the  
succinate/malate CoA ligase alpha subunit family.  
CC -|- PTM: The N-terminus is blocked.  
CC -|- SIMILARITY: In the N-terminal section; belongs to the  
succinate/malate CoA ligase beta subunit family.  
CC -|- SIMILARITY: In the C-terminal section; belongs to the  
succinate/malate CoA ligase alpha subunit family.  
CC -|- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).  
CC -|-  
DR EMBL; AF312052; AAK56081; 1;  
DR EMBL; AF332051; AAK56080; 1;  
DR EMBL; BC056378; AAH56378; 1;  
DR HSSP; P07459; ISCU.  
DR MGD; MGI:103251; AcLY.  
DR InterPro; IPR002020; Citrate synth.  
DR InterPro; IPR003781; CoA binding.  
DR InterPro; IPR005811; CoA Ligase.  
DR InterPro; IPR005810; CoA Lig alpha.  
DR InterPro; IPR005809; CoA binding; 1.  
DR Pfam; PF02629; CoA\_binding; 1.

or send an email to [licensee@isb-sib.ch](mailto:licensee@isb-sib.ch)

or send an email to [licensee@isib-sib.ch](mailto:licensee@isib-sib.ch).

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

EMBL; J05110; AAA74463.1; -.  
 PIR; A35007; A35007.  
 HSSP; P07455; 1SCU.  
 RGD; 2018; Acetyl.  
 InterPro; IPR002020; Citrate synth.  
 InterPro; IPR003781; CoA\_binding.  
 InterPro; IPR003781; CoA\_ligase.  
 InterPro; IPR005810; CoA\_lig\_alpha.  
 InterPro; IPR005809; CoA\_lig\_beta.  
 Pfam; PF02629; CoA\_binding\_1.  
 Pfam; PF00549; Ligase\_CoA\_1.  
 PROSITE; PS01216; SUCCINYL\_COA\_LIG\_1; 1.  
 PROSITE; PS00389; SUCCINYL\_COA\_LIG\_2; 1.  
 PROSITE; PS01217; SUCCINYL\_COA\_LIG\_3; 1.  
 ATP-binding: Direct protein sequencing; Lipid synthesis; Magnesium;  
 Phosphorylation; Transferase.  
 NP\_BIND 700 720 ATP (By similarity).  
 NP\_BIND 751 777 ATP (By similarity).  
 DOMAIN 778 788 CoA-binding (Potential).  
 MOD\_RES 778 788 Phosphoserine (by PKA).  
 ACT\_SITE 454 454 Tele-Phosphohistidine intermediate.  
 METAL 759 759 (By similarity).  
 SEQUENCE 717 717 Magnesium NW;  
 SEQUENCE 1100 AA; 120635 2C6BABC1F3BD2 CR0C4;

```

Query Match      41.6%; Score 52; DB 1; Length 1100;
Best Local Similarity 40.0%; Pred. No. 65;
Matches 10; Conservative 7; Mismatches 8; Indels 0;
1 I V A I A G R I R L M G G D Q F G E L E A S A K 25
| : | : | : | : | : | : | : | : | : | : | : | : |
219 L V S L S T S G L L T D F G G D A A K 943

```

SULT 11  
LY HUMAN

```

--> ACY_HUMAN STANDARD; PRT; 1101 AA.
P53376; Q13037; Q9BR0;
01-OCT-1995 (Rel. 34, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
ATP-citrate synthase (EC 2.3.3.8) (ATP-citrate (pro-S)-lyase)
(Citrate cleavage enzyme).
Name=ACYL;
Name=ACYL;
Homo sapiens (Human). Homo; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Metazoya; Mammalia; Primates; Catarzhini; Hominidae; Homo.
NCBI_TaxID:9606 / [1]
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=92174902; PubMed=1371749;
Elshourbagy N.A., Near J.C., Kmetz P.J., Wells T.N.C., Groot F.,
Saxty B.A., Hughes S.A., Franklin M., Gloger I.S.; "Cloning and expression of the rat ATP-citrate lyase cDNA.", Eur. J. Biochem. 204:491-499 (1992).

```

[2] SEQUENCE FROM N.A.  
TISSUE=Liver;  
MEDLINE=9117247; PubMed=9116495;  
Lord K.A., Wang X.M., Simmons S.J., Bruckner R.C., Loscig J.,  
O'Connor B., Bentley R., Smallwood A., Chadwick C.C., Stevis P.E.,  
Ciccarelli R.B.;  
Variant cDNA sequences of human ATP:citrate lyase: cloning,  
expression, and purification from baculovirus-infected insect cells.";  
Expression Expr. Purif. 9:123-141(1997)

PLoS EXPLORE | DOI:10.1371/journal.pone.024603899  
SEQUENCE FROM N.A.  
TISSUE=Lymph;  
MEDLINE-ID=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shernan C.M., Schuler G.D.  
[3]

RA	Altenschul S.F., Zeeberg B., Buetow K.H., Schaeffer C.F., Bhat N.K.,	
RA	Dopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Marsuska K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J., Mulahy S.J.,	
RA	Raha S.S., Loquillard O.A., Peters G.J., Abramson R.D., Pihl M.C., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarante P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,	
RA	Villalona D.K., Muzny D.M., Scodarigran E.J., Lu X., Gibbs R.A.,	
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shvchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,	
RA	Rodriguez A.C., Grinow J., Schmutz J., Myers R.M.,	
RA	Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smialius D.E.,	
RA	Schnetrich A., Schein J.E., Jones S.J.M., Marra M.A.,	
RA	"Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
CC	-1- FUNCTION: ATP citrate-lyase is the primary enzyme responsible for the synthesis of cytosolic acetyl-CoA in many tissues. Has a central role in de novo lipid synthesis. In nervous tissue it may be involved in the biosynthesis of acetylcholine.	
CC	-1- CATALYTIC ACTIVITY: ADP + phosphoenol + acetyl-CoA + oxaloacetate = ATP + citrate + CoA.	
CC	-1- SUBUNIT: Homotetramer.	
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.	
CC	-1- SIMILARITY: In the N-terminal section: belongs to the succinate/malate CoA ligase beta subunit family. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	
CC	-1- SIMILARITY: In the C-terminal section: belongs to the succinate/malate CoA ligase alpha subunit family.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).	
DR	X64330; QAA45614.1; -.	
DR	U18197; AAB6040.1; -.	
DR	DRB006195; AAH06195.1; -.	
DR	HSSP_P07459; IJKW.	
DR	PIR: S21173; S21173.	
DR	MIM: 108728; -.	
DR	GO: GO:0003346; Citrate lyase complex; TAS.	
DR	GO: GO:0003878; F:ATP citrate synthase activity; TAS.	
DR	GO: GO:0005200; P:ATP catabolism; TAS.	
DR	GO: GO:0006101; P:citrate metabolism; TAS.	
DR	GO: GO:0015936; P:coenzyme A metabolism; TAS.	
DR	InterPro: IPR002020; Citrate synth.	
DR	InterPro: IPR003781; CoA binding.	
DR	InterPro: IPR005811; CoA ligase.	
DR	InterPro: IPR005810; CoA lig alpha.	
DR	Pfam: PF02629; CoA binding.	
DR	Pfam: PF00549; Ligase CoA 1.	
DR	PROSITE: PS00126; SUCCINYL-COA-LIG 1; 1.	
DR	PROSITE: PS00127; SUCCINYL-COA-LIG 3; 1.	
FT	KW ATP-binding; Lipid synthesis; Magnesium; Phosphorylation; Transferase.	
FT	NP-BIND 701 721 ATP (By similarity).	
FT	NP-BIND 752 778 ATP (By similarity).	
FT	DOMAIN 779 789 CoA-binding (Potential).	
FT	MOD-RES 455 455 Phosphoserine (By PKA) (By similarity).	
FT	ACT_SITE 760 760 Tele-phosphonitidine intermediate (By similarity).	
FT	METAL 718 718 Magnesium (By similarity).	
FT	CONFICT 75 75 N > D (In Ref. 1).	
FT	CONFICT 111 111 V > A (In Ref. 1).	
FT	CONFICT 175 175 D > E (In Ref. 3).	



Query Match 40.8%; Score 51; DB 2; Length 1092;  
 Best Local Similarity 36.0%; Pred. No. 92; Gaps 0;  
 Matches 9; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 1 IVAIIAGRLLRMLGDOFNGELEASAK 25  
 ::::: | :| :| :| :| :| :|  
 Db 911 LISSLTSGGLITIGDRGGALDAAK 935

RESULT 14  
 Q70000 PRELIMINARY; PRT; 248 AA.  
 ID Q70000 ;  
 AC Q70000 ;  
 DT 05-JUL-2004 (TREMBurel. 27, Created)  
 DT 05-JUL-2004 (TREMBurel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBurel. 27, Last annotation update)  
 DB Uvrc Protein (Fragment).  
 Name-uvrc:  
 GN Pseudomonas putida.  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 CC NCBI\_TaxID=303; [1]  
 OX RN

SEQUENCE FROM N.A.  
 STRAIN=WCS358;  
 RA Bertani I., Venturi V.;  
 DR Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 EMBL: AJ629219; CAF32984 1; -.  
 InterPro: IPR00943; UVrB/C.  
 InterPro: IPR000305; UVrB\_C.  
 InterPro: IPR000305; UVrC\_N.  
 PFam: PF01541; GIY-YIG; 1.  
 SMART: SM00465; UVr; 1.  
 PROSITE: DSS0151; GIYIC; 1.  
 PROSITE: PS50164; UVr; 1.  
 NON\_TER 248 AA; 248 FT  
 NON\_TER 248 AA; 248 SQ

Query Match 39.2%; Score 49; DB 2; Length 248;  
 Best Local Similarity 39.1%; Pred. No. 43; Gaps 0;  
 Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 VALLAGRLLRMLGDOFNGELEASAK 24  
 ::::: | :| :| :| :| :|  
 Db 195 VMFLERGSQQLGNELNAMEKAA 217

RESULT 15  
 CAF32984 PRELIMINARY; PRT; 248 AA.  
 ID CAF32984 ;  
 AC CAF32984 ;  
 DT 10-MAR-2004 (TREMBurel. 27, Created)  
 DT 10-MAR-2004 (TREMBurel. 27, Last sequence update)  
 DT 10-MAR-2004 (TREMBurel. 27, Last annotation update)  
 DB Uvrc Protein (Fragment).  
 Name-uvrc:  
 GN Pseudomonas putida.  
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 CC NCBI\_TaxID=303; [1]  
 OX RN

SEQUENCE FROM N.A.  
 STRAIN=WCS358;  
 RA Bertani I., Venturi V.;  
 RT "Regulation of the N-acyl homoserine lactone dependent quorum sensing system in rhiosphere Pseudomonas putida WCS358 and cross talk with the stationary phase Rpos sigma factor and the global regulator GacA";  
 RT Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 EMBL: AJ629219; CAF32984 1; -.  
 NON\_TER 248 FT

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 10.6151 Seconds  
(Without alignments)  
181.178 Million cell updates/sec

Title: US-10-092-750-45  
Perfect score: 143  
Sequence: 1 LALAYYSSRQASALKHAETIERGIRQH 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs., 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
 1: /cggn2\_6/podata/1/iaa/5A-COMB.pep:  
 2: /cggn2\_6/podata/1/iaa/5B-COMB.pep:  
 3: /cggn2\_6/podata/1/iaa/6A-COMS.pep:  
 4: /cggn2\_6/podata/1/iaa/6B-COMS.pep:  
 5: /cggn2\_6/podata/1/iaa/PECTUS-COMB.pep:  
 6: /cggn2\_6/podata/1/iaa/backfles1-pep:  
 \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	51	35.7	224	1	US-08-248-466B-8	Sequence 8, Appl1
2	51	35.7	301	4	US-09-489-039A-11121	Sequence 11121, A
3	48.5	33.9	87	3	US-09-134-001C-3005	Sequence 3005, Ap
4	48.5	33.9	224	4	US-09-710-79-3114	Sequence 3114, Ap
5	48.5	33.9	224	4	US-09-710-79-3130	Sequence 31130, Ap
6	48.5	33.9	229	3	US-09-134-001C-5379	Sequence 5379, Ap
7	48.5	33.9	230	3	US-09-134-001C-5907	Sequence 2907, Ap
8	48.5	33.9	230	3	US-09-134-001C-5908	Sequence 2908, Ap
9	48	33.6	463	1	US-07-951-715A-25	Sequence 25, Appl1
10	48	33.6	463	2	US-08-459-448A-25	Sequence 25, Appl1
11	48	33.6	463	3	US-08-459-535A-25	Sequence 25, Appl1
12	48	33.6	463	3	US-08-459-504B-25	Sequence 25, Appl1
13	48	33.6	463	3	US-08-459-444-25	Sequence 0, Appl1
14	48	33.6	463	3	US-09-147-442-25	Sequence 0, Appl1
15	48	33.6	463	4	US-09-988-162-25	Sequence 25, Appl1
16	48	33.6	473	4	US-09-538-992-1353	Sequence 133, Ap
17	46.5	32.9	757	4	US-09-194-597A-2	Sequence 2, Appl1
18	46.5	32.5	197	4	US-10-101-164A-191	Sequence 491, App
19	46	32.2	203	4	US-09-270-767-34950	Sequence 3450, A
20	46	32.2	203	4	US-09-70-767-50167	Sequence 50167, A
21	46	32.2	484	4	US-09-602-787A-26	Sequence 26, Appl1
22	46	32.2	491	3	US-09-362-899-3	Sequence 3, Appl1
23	45	31.5	231	4	US-09-70-167-56309	Sequence 56309, A
24	45	31.5	479	4	US-09-270-167-41666	Sequence 41666, A
25	45	31.5	623	4	US-09-583-110-42926	Sequence 42926, Ap
26	44.5	31.1	230	3	US-09-134-001C-1549	Sequence 5419, Ap
27	44.5	31.1	276	4	US-09-328-352-5243	Sequence 5243, Ap

## ALIGNMENTS

RESULT 1  
US-08-248-466B-8  
; Sequence 8, Application US/08248466B  
; Patent No. 5622182  
GENERAL INFORMATION:  
; APPLICANT: CHOPEN, MARIE-CHRISTINE  
; CLUZEL, PIERRE-JEAN  
; TITLE OF INVENTION: CLONING OF DNA FRAGMENTS ENCODING A  
; MECHANISM FOR RESISTANCE TO BACTERIOPHAGES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 175 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Parent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/248,466B  
; FILING DATE: 24-MAY-1994  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,959  
; FILING DATE: 15-MAR-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: FR 90/11381  
; FILING DATE: 14-SEP-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NO. 56229182 man. F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 384-033-0 PCT FMC CIP  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 224 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-248-466B-8  
Query Match 35.7%; Score 51; DB 1; Length 224;

Best Local Similarity 33.3%; Pred. No. 5.8;  
 Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

3 LAYSSRQYASALKHAEIIEERGIRHQ 29  
 186 LSYVSSKQEAVSKRNAYKLVKAKIEQY 212

SULT 2  
 -03-489-039A-11121  
 Sequence 11121, Application US/09489039A  
 Patent No. 6110836

GENERAL INFORMATION:  
 APPLICANT: Gary Breton et. al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709-2004001  
 CURRENT APPLICATION NUMBER: US/09/489-039A  
 CURRENT FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: US 60/117,747  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 14342  
 SEQ ID NO 11121  
 LENGTH: 301  
 TYPE: PRT  
 ORGANISM: Klebsiella pneumoniae  
 -03-489-039A-11121

Query Match	Score	DB	Length
Best Local Similarity	35.7%	4	301
Matches	40.7%	No.	8.1
11; Conservative	9;	Mismatches	
		Indels	2;
		Gaps	
1 LALAYYSSROYASALKHIAETBERGIR 27	:	:	:
9 ILLTYSNRQ -AALDVVAEIEQKGVK 103	:	:	:

SUIT 3  
c-09-134-001C-3005  
Sequence 3005, Application US/09134001C  
Patient No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

FILE REFERENCE: GIC-00  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
NUMBER OF SEQ ID NOS: 5674

Query	Match	33.3%	Score	48.5	DB	3	Length	87
BEST Local	Similarity	40.0%	Pred. No.	4.6				
matches	12;	Conservative	B; Mismatches	8;	Indels	7	Gaps	2;
09-134-001C-3005								

1 LALAYSSROYASALKHAETI-ERGIROH 29

22 VAVGY - LRYALSYRDISEILLRERGV  
S0014 0-9-710-279-3114  
Sequence 1114, Application US/09710279  
Patent No. 67034392

```

; APPLICANT: KIMMELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P13480TS
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIORITY FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3114
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; US-09-710-279-3114

Query Match          33.9% Score 48.5; DB 4; Length 224;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 12; Conservative 8; Mismatches 7; Indels 3; Gaps 2;
Db      1  LALAYISSRQVASALKHIAEII-ERGIROH 29
          :|:||:||:||:||:||:||:|
Db      16  VAVGYY--LRLVALSYRDISEIILRRGVNVH 43

```

US-09-710-079-3130  
; Sequence 3130, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P134801S  
; COUNTRY: UNITED STATES OF AMERICA  
; STATEMENT OF INVENTION:

CURRENT FILING DATE: 2000-11-09  
PRIORITY APPLICATION NUMBER: 60/164,258  
PRIORITY FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 3130  
LENGTH: 224  
TYPE: PRT  
ORGANISM: [artificial] Semience

```

; FAILURE;
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-3130

Query Match          33.9%; Score 48.5; DB 4; Length 224;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 12; Conservative 8; Mismatches 7; Indels 3; Gaps 2;

```

**RESULT 6**  
 US-09-134-001C-3379  
 ; Sequence 5379 ; Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al

Qy	1	LALAYISSROYASALGHIAELI-EIGGIRQH	29
	:	-  :   :   :   :   :	
Db	16	VAVGYY - LRYALSYRDIESEILRERGVNVH	43

TITLE OF INVENTION: NUCLEIC ACID AND AMINO  
ACID ANALOGUE  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSIS  
FILE REFERENCE: GTC-1007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08

PRIOR FILING DATE: 1987-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5379  
LENGTH: 229  
TYPE: PRF  
ORGANISM: *Staphylococcus epizooticus*  
US 09-134-005C 5379

RESULT 7  
US 09-134-001C-2907  
Sequence 2907, Application US/09134 001C  
Patient No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,364  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO: 2907  
LENGTH: 230  
TYPE: PRT

ORGANISM: *Staphylococcus epidermidis*

Query Match      Score 4  
 Best Local Similarity      33.9%;  
 Pred. N  
 Matches 12;      Conservative 40.0%;  
 Mism.

RESULT<sup>8</sup>  
US-09-134-001C-2908  
Sequence 2908, Application US/09134001C  
Patient No. 6384370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,364  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2908  
LENGTH: 230  
TYPE: PCT  
ORGANISM: *Staphylococcus epidermidis*  
US-09-134-001C-2908  
Query Match 33.9% Score 48.5% DB 3; Length 230;  
Test Local Similarity 40.0% Pre. No. 14; Mismatches 8; Indels 3; Gaps 2;

QY 1 LALAYSSROYASALKHAEII-ERGIROH 29  
           : : | : | : | : | : | : |  
 DB 22 VAVGYY--LRYALSYRDISEILRERGVNVH 49

RESULT 9  
US-07-951-715A-25  
; Sequence 25, Application US/07951715A

GENERAL INFORMATION:  
APPLICANT: Kozai, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.

APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lytle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttle, Janet L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
FUNCTION

NAME OF INVENTOR: INSPECTORATE, A DIVISION, TN MATTER

NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:

PATENT APPLICATION NUMBER: 08/012345, 11/354  
FILING DATE: 25-SEP-1992  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772, 027  
FILING DATE: 04-07-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32, 943  
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-1615  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 25:

```

; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
;   NAME/KEY: Protein
;   LOCATION: 1..463
;   OTHER INFORMATION: /note="protein sequence for
;   OTHER INFORMATION: soybean CDPK as shown in Figure 34."
;   OTHER INFORMATION: US-07-951-715A-25

```

Best Local Similarity 47.4% Pred. No. 387

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0; OTHER INFORMATION: soybean CDPK as shown in Figure 34."

Qy 5 YYSSRQYASAFKHIAELIE 23  
Db 108 HYSERQARLKITIVVE 126

US-08-459-448A-25

Sequence 25, Application US/08459448A  
Patent No. 5859336

GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Sutie, Janet L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5859336artis Corporation  
STREET: Patent & Trademark Dept., 520 White Plains  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,448A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/459,448A  
FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CCC 1577/CIP/DIV4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 463 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..463  
OTHER INFORMATION: /note= "protein sequence for

OTHER INFORMATION: US-08-459-448A-25

Query Match 33.6%; Score 48; DB 2; Length 463;  
Best Local Similarity 47.4%; Pred. No. 38;  
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YYSSRQYASAFKHIAELIE 23  
Db 108 HYSERQARLKITIVVE 126

RESULT 11  
US-08-459-505A-25  
Sequence 25, Application US/08459595A

Patent No. 6018104

GENERAL INFORMATION:

APPLICANT: Kozel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Sutie, Janet L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6018104artis Corporation  
STREET: Rd. POB 2005  
CITY: White Plains  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,595A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/459,595A  
FILING DATE: 04-OCT-1991  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CCC 1577/CIP/DIV4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 463 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..463  
OTHER INFORMATION: /note= "protein sequence for

RESULT 12  
US-08-459-505A-25  
Sequence 25, Application US/08459595A

Patent No. 6018104

GENERAL INFORMATION:

APPLICANT: Kozel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Sutie, Janet L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6018104artis Corporation  
STREET: Rd. POB 2005  
CITY: White Plains  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,595A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/459,595A  
FILING DATE: 04-OCT-1991  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CCC 1577/CIP/DIV4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 463 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..463  
OTHER INFORMATION: /note= "protein sequence for

HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..463  
 OTHER INFORMATION: /note= "protein sequence for soybean CDPK as shown in Figure 34."  
 US-08-459-595A-25

Query Match Score 48; DB 3; length 463;  
 Best Local Similarity 47.4%; Pred. No. 38;  
 Matches 9; Conservative 4; Mismatches 6;  
 Indels 0; Gaps 0;

Qy 5 YYSSROYAYAKKIIAHIE 23  
 Db 108 HYSERQARLKITIVVVV 126

---

RESULT 12  
 Sequence 25, Application US/08459504B  
 Patent No. 6075185

GENERAL INFORMATION:  
 APPLICANT: Koziel, Michael G.  
 APPLICANT: Desai, Nalini M.  
 APPLICANT: Lewis, Kelly S.  
 APPLICANT: Kramer, Vance C.  
 APPLICANT: Warren, Gregory W.  
 APPLICANT: Brobla, Stephen V.  
 APPLICANT: Crossland, Lyle D.  
 APPLICANT: Wright, Martha S.  
 APPLICANT: Merlin, Ellis J.  
 APPLICANT: Launis, Karen L.  
 APPLICANT: Rothstein, Steven J.  
 APPLICANT: Bowman, Cindy G.  
 APPLICANT: Dawson, John L.  
 APPLICANT: Dunder, Erik M.  
 APPLICANT: Pace, Gary M.  
 APPLICANT: Sutcliffe, Janet J.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
 TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
 NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6075185artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/459, 504B  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/459, 595  
 FILING DATE: 02-JUN-1995  
 APPLICATION NUMBER: US 07/951, 715  
 FILING DATE: 25-SEP-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/772, 027  
 FILING DATE: 04-OCT-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38, 241  
 REFERENCE/DOCKET NUMBER: S-18805/P1/GCG1577/CIP/DIV6  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919)541-8587  
 TELEFAX: (919)541-8689  
 INFORMATION FOR SEQ ID NO: 25:

---

SEQUENCE CHARACTERISTICS:  
 LENGTH: 463 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..463  
 OTHER INFORMATION: /note= "protein sequence for soybean CDPK as shown in Figure 34."  
 US-08-459-595A-25

Query Match Score 48; DB 3; length 463;  
 Best Local Similarity 47.4%; Pred. No. 38;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YYSSROYAYAKKIIAHIE 23  
 Db 108 HYSERQARLKITIVVVV 126

---

RESULT 13  
 Sequence 0, Application US/08459444A  
 Patent No. 6121014

GENERAL INFORMATION:  
 APPLICANT: Koziel, Michael G.  
 APPLICANT: Desai, Nalini M.  
 APPLICANT: Lewis, Kelly S.  
 APPLICANT: Kramer, Vance C.  
 APPLICANT: Warren, Gregory W.  
 APPLICANT: Evola, Stephen V.  
 APPLICANT: Crossland, Lyle D.  
 APPLICANT: Merlin, Ellis J.  
 APPLICANT: Launis, Karen L.

TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED  
 NUCLEIC ACID CODING SEQUENCE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/459, 444A  
 FILING DATE: 02-Jun-1995  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/951, 715  
 FILING DATE: 25-SEP-1992  
 APPLICATION NUMBER: US 07/772, 027  
 FILING DATE: 04-OCT-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38, 241  
 REFERENCE/DOCKET NUMBER: S-18805/P1/GCG1577/CIP/DIV6  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919)541-8587  
 TELEFAX: (919)541-8689  
 INFORMATION FOR SEQ ID NO: /note= "protein sequence for soybean CDPK as shown in Figure 34."  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 463 amino acids

TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..463  
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
 US-08-459-144-25

Query Match 33.6%; Score 48; DB 3; Length 463;  
 Best Local Similarity 47.4%; Pred. No. 38;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YYSSRQVASALKHIAELIE 23  
 Db 108 HYSERQAARLIKATIVEVE 126

RESULT 14  
 US-09-547-422-25  
 Sequence 0, Application US/09547422  
 Patent No. 6320100  
 GENERAL INFORMATION:  
 APPLICANT: Koziel, Michael G.  
 Desai, Nalini M.  
 Lewis, Kelly S.  
 Kramer, Vance C.  
 Warren, Gregory W.  
 Evola, Stephen V.  
 Crossland, Lytle D.  
 Wright, Martha S.  
 Merlin, Ellis J.  
 Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/547,422  
 FILING DATE: 11-APR-2000  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/459,595  
 FILING DATE: 02-JUN-1995  
 APPLICATION NUMBER: US 07/951,715  
 FILING DATE: 25-SEP-1992  
 APPLICATION NUMBER: US 07/772,027  
 FILING DATE: 04-OCT-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38  
 REFERENCE/DOCKET NUMBER: S-18805H

TELEPHONE: (919)541-8587  
 TELEFAX: (919)541-8589

INFORMATION FOR SEQ ID NO: /noe= "protein sequence for soybean CDPK as shown in Figure 34."  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 463 amino acids  
 TYPE: amino acid

STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..463  
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
 US-09-547-422-25

Query Match 33.6%; Score 48; DB 3; Length 463;  
 Best Local Similarity 47.4%; Pred. No. 38;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YYSSRQVASALKHIAELIE 23  
 Db 108 HYSERQAARLIKATIVEVE 126

RESULT 15  
 US-09-588-462-25  
 Sequence 25, Application US/0988462  
 Patent No. 6720488  
 GENERAL INFORMATION:  
 APPLICANT: Koziel, Michael G.  
 Desai, Nalini M.  
 Lewis, Kelly S.  
 Kramer, Vance C.  
 Warren, Gregory W.  
 Evola, Stephen V.  
 Crossland, Lytle D.  
 Wright, Martha S.  
 Merlin, Ellis J.  
 Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Syngenta Biotechnology, Inc.  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/588-462  
 FILING DATE: 20-NOV-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38  
 REFERENCE/DOCKET NUMBER: S-18805I

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919)541-8587  
 TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 463 amino acids  
 TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..463
; OTHER INFORMATION: /note= "protein sequence for
; soybean CDPK as shown in Figure 34."
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-988-462-25

Query Match          Score 48; DB 4; Length 463;
Best Local Similarity 33.6%; Pred. No. 38;
Matches 9; Conservative 4%; Mismatches 6; Indels 0; Gaps 0;
;                                     5 YYSSRQYASALKHIAELIE 23
;                                     :|:||| | :| |::|:
Db                               108 HYSERQARLKITIVVEVE 126

```

Search completed: November 10, 2004, 14:55:34  
Job Time: 11:56:51 since

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 33.5415 Seconds  
(without alignments)  
305.399 Million cell updates/sec

Title: US-10-092-750-45  
Perfect score: 143  
Sequence: 1 LAIAYYSSROYASALKHAETIERGIRQH 29

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA;\*

1: /cggn2\_6/prodata/1/pubpaas/US07\_PUBCOMB.pep;\*

2: /cggn2\_6/prodata/1/pubpaas/PCT\_NEW\_PUB.pep;\*

3: /cggn2\_6/prodata/1/pubpaas/US05\_NEW\_PUB.pep;\*

4: /cggn2\_6/prodata/1/pubpaas/US06\_PUBCOMB.pep;\*

5: /cggn2\_6/prodata/1/pubpaas/US07\_NEW\_PUB.pep;\*

6: /cggn2\_6/prodata/1/pubpaas/PCTOS\_PUBCOMB.pep;\*

7: /cggn2\_6/prodata/1/pubpaas/US08\_NEW\_PUB.pep;\*

8: /cggn2\_6/prodata/1/pubpaas/US08\_PUBCOMB.pep;\*

9: /cggn2\_6/prodata/1/pubpaas/US09\_PUBCOMB.pep;\*

10: /cggn2\_6/prodata/1/pubpaas/US09B\_PUBCOMB.pep;\*

11: /cgpn2\_6/prodata/1/pubpaas/US09C\_PUBCOMB.pep;\*

12: /cgpn2\_6/prodata/1/pubpaas/US09\_NEW\_PUB.pep;\*

13: /cgpn2\_6/prodata/1/pubpaas/US10A\_PUBCOMB.pep;\*

14: /cgpn2\_6/prodata/1/pubpaas/US10B\_PUBCOMB.pep;\*

15: /cgpn2\_6/prodata/1/pubpaas/US10C\_PUBCOMB.pep;\*

16: /cgpn2\_6/prodata/1/pubpaas/US10D\_PUBCOMB.pep;\*

17: /cgpn2\_6/prodata/1/pubpaas/US11\_NEW\_PUB.pep;\*

18: /cgpn2\_6/prodata/1/pubpaas/US11\_NEW\_PUB.pep;\*

19: /cgpn2\_6/prodata/1/pubpaas/US60\_NEW\_PUB.pep;\*

20: /cgpn2\_6/prodata/1/pubpaas/US60\_PUBCOMB.pep;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1  
US-10-092-750-45  
Sequence 45, Application US/10092750  
; Publication No. US2003003215711  
; GENERAL INFORMATION:  
; APPLICANT: Hammond, Philip W.  
; APPLICANT: Alpin, Julia  
; APPLICANT: Wright, Martin C.  
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1  
; CURRENT APPLICATION NUMBER: US/10/092,750  
; CURRENT FILING DATE: 2002-03-07  
; PRIORITY APPLICATION NUMBER: US 60/274,526  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 45  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-750-45

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	29	US-10-092-750-45	Sequence 45, Appl
2	143	100.0	628	14	Sequence 3203, Ap
3	143	100.0	638	14	Sequence 2059, Ap
4	143	100.0	638	17	Sequence 55, Appl
5	143	100.0	665	15	Sequence 74, Appl
6	54	37.8	131	14	Sequence 104, App
7	52.5	36.7	109	15	Sequence 27504,
8	51	35.7	238	16	Sequence 32647, A
9	51	35.7	382	17	Sequence 281833, A
10	51	35.7	398	15	Sequence 53195, A
11	51	35.7	398	15	Sequence 64162, A
12	51	35.7	398	15	Sequence 72211, A
13	50	35.0	555	15	Sequence 60515, A

Query Match 100.0%; Score 143; DB 14; Length 29;  
Best Local Similarity 100.0%; Pred. No. 8.6e-14;  
Matches 29; Conservative 0; Mismatches 0; Index 0; Gaps 0;  
Qy 1 LALAYYSSROYASALKHAETIERGIRQH 29  
Db 1 LALAYYSSROYASALKHAETIERGIRQH 29  
RESULT 2  
US-10-094-749-3203  
Sequence 3203, Application US/10094749  
; Publication No. US20030219749.A1  
; GENERAL INFORMATION:

PRIOR FILING DATE: 2001-09-14  
 NUMBER OF SEQ ID NOS: 3381  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 2099  
 LENGTH: 638  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-094-749-2099

Query Match 100.0%; Score 143; DB 14; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

RESULT 4  
 US-10-786-720-55  
 Sequence 55, Application US/10786720  
 Publication No. US2004019181A1  
 GENERAL INFORMATION:  
 APPLICANT: Nyeth  
 APPLICANT: O'Toole, Margot  
 APPLICANT: Liu, Wei  
 APPLICANT: Liou, Wei  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING AUTOIMMUNE  
 TITLE OF INVENTION: DISEASES  
 FILE REFERENCE: 031896-02300 (AM101311)  
 CURRENT APPLICATION NUMBER: US10/786,720  
 NUMBER OF SEQ ID NOS: 21135  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 55  
 LENGTH: 628  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-094-749-3203

Query Match 100.0%; Score 143; DB 14; Length 628;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

RESULT 5  
 US-10-210-172-74  
 Sequence 74, Application US/10210172  
 Publication No. US20040043928A1  
 GENERAL INFORMATION:  
 APPLICANT: Kekuda, Ramesh  
 APPLICANT: Miller, Charles  
 APPLICANT: Pattarajan, Meera  
 APPLICANT: Pena, Carol  
 APPLICANT: Rieger, Daniel  
 APPLICANT: Shirkets, Richard  
 APPLICANT: Zerhusen, Bryan  
 APPLICANT: Li, Li  
 APPLICANT: Ji, Weizhen  
 APPLICANT: Padigaru, Muralidhara  
 APPLICANT: Casman, Shacie  
 APPLICANT: Voss, Edward  
 APPLICANT: Boldog, Ferenc  
 APPLICANT: Gorman, Linda  
 APPLICANT: Leite, Mario  
 APPLICANT: Vernet, Corinne  
 APPLICANT: Anderson, David  
 APPLICANT: Guo, Xiaojia  
 APPLICANT: Zhong, Mei  
 APPLICANT: Gerlach, Valerie

FILE REFERENCE: 084335/0160  
 CURRENT FILING DATE: 2002-03-12  
 PRIOR APPLICATION NUMBER: 60/0350,435  
 PRIOR FILING DATE: 2002-01-24  
 PRIOR APPLICATION NUMBER: JP 2001-328381  
 CURRENT FILING DATE: 2001-09-14  
 PRIOR APPLICATION NUMBER: 60/3350,435  
 NUMBER OF SEQ ID NOS: 3381  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 3203  
 LENGTH: 628  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-094-749-3203

Query Match 100.0%; Score 143; DB 14; Length 628;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

RESULT 3  
 US-10-094-749-2099  
 Sequence 2099, Application US/10094749  
 Publication No. US2003021971A1  
 GENERAL INFORMATION:  
 APPLICANT: SUGIYAMA, TOMOYASU  
 APPLICANT: OTSUKI, TETSUJI  
 APPLICANT: WAKAMATSU, AI  
 APPLICANT: SATO, HIRROYUKI  
 APPLICANT: ISHII, SHIZUKO  
 APPLICANT: YAMAMOTO, JUN-ICHI  
 APPLICANT: ISONO, YUTIRO  
 APPLICANT: HIO, YURI  
 APPLICANT: OTSUKA, KAORU  
 APPLICANT: NAGAHARI, KENJI  
 APPLICANT: MASUO, TSUYUKI  
 TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
 FILE REFERENCE: 084335/0160  
 CURRENT FILING DATE: 2002-03-12  
 PRIOR APPLICATION NUMBER: 60/0350,435  
 PRIOR FILING DATE: 2002-01-24  
 PRIOR APPLICATION NUMBER: JP 2001-328381

Query Match 100.0%; Score 143; DB 14; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Query Match 100.0%; Score 143; DB 14; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Query Match 100.0%; Score 143; DB 14; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Query Match 100.0%; Score 143; DB 14; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
 Matches 29; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

APPLICANT: Hjalt, Tord  
 APPLICANT: Rastelli, Luca  
 APPLICANT: Spytek, Kimberly  
 APPLICANT: Edinger, Shlomit  
 APPLICANT: Malynak, Uriel  
 APPLICANT: Ellerman, Karen  
 APPLICANT: Macdougal, John  
 APPLICANT: Stone, David  
 APPLICANT: Alsobrook II, John  
 APPLICANT: Lepley, Denise et al.  
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 FILE REFERENCE: 21402-416 A  
 CURRENT APPLICATION NUMBER: US/10/210,172  
 CURRENT FILING DATE: 2001-08-01  
 PRIOR APPLICATION NUMBER: 60/309,501  
 PRIOR FILING DATE: 2001-08-02  
 PRIOR APPLICATION NUMBER: 60/323,994  
 PRIOR FILING DATE: 2001-09-21  
 PRIOR APPLICATION NUMBER: 60/373,814  
 PRIOR FILING DATE: 2002-04-19  
 PRIOR APPLICATION NUMBER: 60/310,291  
 PRIOR FILING DATE: 2001-08-03  
 PRIOR APPLICATION NUMBER: 60/310,951  
 PRIOR FILING DATE: 2001-08-08  
 PRIOR APPLICATION NUMBER: 60/310,544  
 PRIOR FILING DATE: 2001-08-07  
 PRIOR APPLICATION NUMBER: 60/311,292  
 PRIOR FILING DATE: 2001-08-09  
 PRIOR APPLICATION NUMBER: 60/311,979  
 PRIOR FILING DATE: 2001-08-13  
 PRIOR APPLICATION NUMBER: 60/313,201  
 PRIOR FILING DATE: 2001-08-17  
 PRIOR APPLICATION NUMBER: 60/312,892  
 PRIOR FILING DATE: 2001-08-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 327  
 SEQ ID NO: 74  
 LENGTH: 665  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-210-172-74

Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 29;保守型 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LALAYSSROYASALKHIAETTERGIRQH 29  
 Db 194 LALAYSSROYASALKHIAETTERGIRQH 222

RESULT 6  
 US-10-369-493-106  
 ; Sequence 106, Application US/10369493  
 ; Publication No. US2003233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10/52052 B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/3360,039  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO: 106  
 ; LENGTH: 131  
 ; TYPE: PRT

Query Match Score 54%; Best Local Similarity 22.4%; Matches 13;保守型 7; Mismatches 8; Indels 30; Gaps 1;

Qy 1 LALAYSSROYASALKHIAETTERGIRQH 28  
 Db 49 LGLEYFKVDENAIKHLERYDELQDEGAATYRLCYEEGEFFAKAEVLEEGIRQ 106

RESULT 7  
 US-10-425-115-276504  
 ; Sequence 276504, Application US/10425115  
 ; Publication No. US2004014272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21/53222 B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 3693326  
 ; SEQ ID NO: 276504  
 ; LENGTH: 109  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)-(109)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_183754C.1.pep  
 US-10-425-115-276504

Query Match Score 52.5%; Best Local Similarity 35.7%; Matches 10;保守型 7; Mismatches 10; Indels 1; Gaps 1;

Qy 3 LAYYSROYASALKHIAETTERGIRQH 29  
 Db 68 IRHHHMQRVPAHSMSMVPHGFLRH 95

RESULT 8  
 US-10-767-701-32647  
 ; Sequence 32647, Application US/10767701  
 ; Publication No. US20040172684A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21/5353 B  
 ; CURRENT APPLICATION NUMBER: US/10/767,701  
 ; NUMBER OF SEQ ID NOS: 63128  
 ; SEQ ID NO: 32647  
 ; LENGTH: 238  
 ; TYPE: PRT  
 ; ORGANISM: Sorghum bicolor  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1459\_1.pep  
 US-10-767-701-32647

Query Match Score 51%; Best Local Similarity 50.0%; Matches 10;保守型 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 YSSRQYASALKHIAELIERG 25  
 Db 171 YSNPGVLDALKHITDLKEEG 190

RESULT 9  
 US-10-425-115-281833 ; Sequence 281833, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO: 281833  
 ; LENGTH: 382  
 ; TYPE: PRT  
 ; FEATURE: Zea mays  
 ; OTHER INFORMATION: Clone ID: MRT4577\_20131C.1.pep

US-10-425-115-281833  
 Query Match Score 51; DB 17; Length 382;  
 Best Local Similarity 50%; Pred. No. 51;  
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 YSSRQYASALKHIAELIERG 25  
 Db 170 YSNPGVLDALKHITDLKEEG 189

RESULT 10  
 US-10-425-114-53395 ; Sequence 53395, Application US/10425114  
 ; Publication No. US2004034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO: 53395  
 ; LENGTH: 398  
 ; TYPE: PRT  
 ; FEATURE: Zea mays  
 ; OTHER INFORMATION: Clone ID: 700451509\_FLI.pep

US-10-425-114-53395  
 Query Match Score 51; DB 15; Length 398;  
 Best Local Similarity 50%; Pred. No. 53;  
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 YSSRQYASALKHIAELIERG 25  
 Db 186 YSNPGVLDALKHITDLKEEG 205

RESULT 11  
 US-10-425-114-64162 ; Sequence 64162, Application US/10425114  
 ; Publication No. US2004034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10425114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO: 64162  
 ; LENGTH: 398  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3592-097-D6\_FLI.pep

US-10-425-114-64162  
 Query Match Score 51; DB 15; Length 398;  
 Best Local Similarity 50%; Pred. No. 53;  
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 YSSRQYASALKHIAELIERG 25  
 Db 186 YSNPGVLDALKHITDLKEEG 205

RESULT 12  
 US-10-425-114-72211 ; Sequence 72211, Application US/10425114  
 ; Publication No. US2004034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10425114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO: 72211  
 ; LENGTH: 398  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700428946\_FLI.pep

US-10-425-114-72211  
 Query Match Score 51; DB 15; Length 398;  
 Best Local Similarity 50%; Pred. No. 53;  
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 YSSRQYASALKHIAELIERG 25  
 Db 186 YSNPGVLDALKHITDLKEEG 205

RESULT 13  
 US-10-28-1122A-60515 ; Sequence 60515, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu

FILE REFERENCE: ELITRA.034A ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

CURRENT APPLICATION NUMBER: US/10/282,122A ; CURRENT FILING DATE: 2003-02-20 ; PRIORITY NUMBER: 60/191,078

PRIOR APPLICATION NUMBER: 60/205,848 ; PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/207,727 ; PRIOR FILING DATE: 2000-05-16

PRIOR APPLICATION NUMBER: 60/230,335 ; PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347 ; PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578 ; PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625 ; PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931 ; PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636 ; PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308 ; PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614 ; SEQ ID NO: 60315 ; LENGTH: 555

TYPE: PRT ; ORGANISM: Listeria monocytogenes ; US-10-282-122A-60515

Query Match 35.0% ; Score 50; DB 15; Length 555; Best Local Similarity 62.5%; Pred. No. 1.1e+02; Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 YASALKHAETIERGI 26  
Db 175 YASDLISHIAEFGEKGV 190

RESULT 14 US-10-369-493-22933 ; Sequence 22933, Application US/10369493 ; PUBLICATION NO. US2003233675A1 ; GENERAL INFORMATION: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES ; FILE REFERENCE: 3-10(52052)B ; CURRENT FILING DATE: 2003-02-28 ; PRIOR APPLICATION NUMBER: US/10/369,493 ; NUMBER OF SEQ ID NOS: 47374 ; LENGTH: 272

Query Match 34.3% ; Score 49; DB 15; Length 380; Best Local Similarity 63.2%; Pred. No. 99; Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 1;

Qy 1 LALAYSSROYASALKHAEIERGI 26  
Db 24 VGLGTAIRDYSSALESYAIEERGI 49

RESULT 15 US-10-282-122A-49861 ; Sequence 49861, Application US/10282122A ; PUBLICATION NO. US20040029129A1 ; GENERAL INFORMATION: ; APPLICANT: Wang, Liangsu ; APPLICANT: Zamudio, Carlos ; APPLICANT: Malone, Cheryl ; APPLICANT: Haselbeck, Robert ; APPLICANT: Ohlsen, Kari ; APPLICANT: Zyskind, Judithn ; APPLICANT: Wall, Daniel ; APPLICANT: Trawick, John ; APPLICANT: Carr, Grant ; APPLICANT: Yamamoto, Robert ; APPLICANT: Forsyth, R. ; APPLICANT: Xu, H. ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms ; FILE REFERENCE: ELITRA.034A ; CURRENT APPLICATION NUMBER: US/10/282,122A ; CURRENT FILING DATE: 2003-02-20 ; PRIOR APPLICATION NUMBER: 60/191,078 ; PRIOR FILING DATE: 2000-03-21 ; PRIOR APPLICATION NUMBER: 60/206,848 ; PRIOR FILING DATE: 2000-05-23 ; PRIOR APPLICATION NUMBER: 60/207,727 ; PRIOR FILING DATE: 2000-05-16 ; PRIOR APPLICATION NUMBER: 60/230,335 ; PRIOR FILING DATE: 2000-09-06 ; PRIOR APPLICATION NUMBER: 60/230,347 ; PRIOR FILING DATE: 2000-09-09 ; PRIOR APPLICATION NUMBER: 60/242,578 ; PRIOR FILING DATE: 2000-10-23 ; PRIOR APPLICATION NUMBER: 60/253,625 ; PRIOR FILING DATE: 2000-11-27 ; PRIOR APPLICATION NUMBER: 60/257,931 ; PRIOR FILING DATE: 2000-09-09 ; PRIOR APPLICATION NUMBER: 60/267,636 ; PRIOR FILING DATE: 2000-12-22 ; PRIOR APPLICATION NUMBER: 60/269,308 ; PRIOR FILING DATE: 2000-02-09 ; PRIOR APPLICATION NUMBER: 60/270,727 ; PRIOR FILING DATE: 2000-05-25 ; PRIOR APPLICATION NUMBER: 60/273,035 ; PRIOR FILING DATE: 2000-09-06 ; PRIOR APPLICATION NUMBER: 60/280,347 ; PRIOR FILING DATE: 2000-09-09 ; PRIOR APPLICATION NUMBER: 60/242,578 ; PRIOR FILING DATE: 2000-10-23 ; PRIOR APPLICATION NUMBER: 60/253,625 ; PRIOR FILING DATE: 2000-11-27 ; PRIOR APPLICATION NUMBER: 60/257,931 ; PRIOR FILING DATE: 2000-02-16 ; Remaining Prior Application data removed - See File Wrapper or PALM. ; NUMBER OF SEQ ID NOS: 78614 ; SEQ ID NO: 49861 ; SOFTWARE: PatentIn version 3.1 ; TYPE: PRT ; ORGANISM: Burkholderia fungorum ; US-10-282-122A-49861

Query Match 34.3% ; Score 49; DB 15; Length 380; Best Local Similarity 63.2%; Pred. No. 99; Matches 12; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 9 ROYASALKHAEIERGI 27  
Db 56 RGYVEALKHIEPIL-GIR 72

SEQ ID NO: 22933 ; LENGTH: 272

Search completed: November 11, 2004, 02:43:01

Fri Nov 12 14:55:29 2004

us-10-092-750-45.rapb

Page 6

Job time : 34.5915 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

Om protein - protein search, using sw model  
 Run on: November 10, 2004, 13:40:53 ; Search time 7.11321 Seconds  
 (without alignments)  
 392.268 Million cell updates/sec

Title: US-10-092-750-45  
 Perfect score: 143  
 Sequence: 1 LALAYYSSROYASALKHAETIERGIRQH 29

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR 79:  
 1: Pir1:  
 2: Pir2:  
 3: Pir3:  
 4: Pir4:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	76	53.1	418	2	F87767		protein F54C1.5 [1
2	54	37.8	131	2	F70422		hypothetical prote
3	53	37.1	368	2	A82249		response regulator
4	52	36.4	1227	2	T49963		hypothetical prote
5	51	35.7	280	2	C64471		hypothetical prote
6	51	35.7	381	2	S37170		repB protein - Lac
7	51	35.7	388	2	S40057		repA protein - Lac
8	51	35.7	403	2	A28252		Na+Pi-cotransport
9	51	35.7	550	2	B97629		hypothetical prote
10	51	35.7	555	2	B97629		protein FS01.2 [1
11	51	35.7	572	2	H90458		hypothetical prote
12	50.5	35.3	316	2	D69822		ABC transporter (A
13	50	35.0	305	2	B93861		hypothetical prote
14	50	35.0	338	2	AB1254		hypothetical prote
15	50	35.0	555	2	AH1616		hypothetical prote
16	35.0	35.0	555	2	C72498		probable stress pr
17	49	34.3	272	2	A95861		hypothetical prote
18	49	34.3	547	2	B60634		tetratricopeptide
19	49	34.3	596	2	G75457		hypothetical prote
20	49	34.3	672	2	AC0422		competence protein
21	49	34.3	691	2	S39867		competence protein
22	49	34.3	691	2	G81167		conformational source
23	49	34.3	691	2	B81937		probable n-termina
24	49	34.3	727	2	T41735		probable transposa
25	48.5	33.9	221	2	A60634		probable transposa
26	48.5	33.9	224	2	C60634		probable transposa
27	48.5	33.9	224	2	A60634		heterodisulfide re
28	48	33.6	145	2	H69051		conserved hyProphet
29	48	33.6	276	2	AH0460		

## ALIGNMENTS

## RESULT 1

F87767 protein F54C1.5 [imported] - Caenorhabditis elegans

CISpecies: Caenorhabditis elegans

CDate: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

CAccession: F87767

Rianyamou, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A.Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A.Reference number: A75000; PMID:99069611;

A.Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/

A.Accession: F87767

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-418 &lt;STO&gt;

A.Cross-references: GB:chr\_I; PIDN:AAB93327.1; PID:q1825574; GSPDB:GN00019; CESS:F54C1.5

C.Genetics:

A.Gene: F54C1.5

A.Map position: 1

Query Match Score 76; DB 2%; Length 416;

Best Local Similarity 51.1%;保守性 Matches 15; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 LALAYYSSROYASALKHAETIERGIRQH 29

Db 159 IALCHYRRGDYDSALKLISEINRGV7RDH 187

## RESULT 2

F70422 hypothetical protein aq\_1409 - Aquifex aeolicus

CISpecies: Aquifex aeolicus

CDate: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 16-Aug-2004

CAccession: F70422

R.J.Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

V.

Nature 392, 353-358, 1998

A.Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A.Reference number: A70300; PMID:9819666;

A.Accession: F70422

A.Status: Preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-131 &lt;AOF&gt;

A.Cross-references: UNIPROT:O57408; GB:AE000739; PIDN:AC07377.1; PID:9298.

A.Experimental source: strain VFS

A.Genetics:

A.Gene: aq\_1409

C.Superfamily: tetratricopeptide repeat homology &lt;TT1&gt;

F13-76/Domain: tetratricopeptide repeat homology &lt;TT2&gt;

F17-110/Domain: tetratricopeptide repeat homology &lt;TT2&gt;

Query Match 37.8%; Score 54; DB 2; Length 131;  
Best Local Similarity 22.4%; Pred. No. 1.7; Matches 13; Conservative 7; Nismatches 8; Indels 30; Gaps 1;

Qy 1 LALAYSSRQYASALKHII-----ABIEBIRGIRQ 28  
Db 49 IGLLEYFKVQDXENAIKHLERYLEQDEGGAAVRTLAKCYELGBFQKAIEVLEBIRGIRQ 106

**RESULT 3**  
A82249 response regulator VC1050 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
C;Species: *Vibrio cholerae*  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: A82249  
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Ginn, M.L.; Dodson, R.J./  
charlson, D.; Embley, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A;Reference number: A82055; PMID:20406833; PIDN:10952201  
A;Accession: A82249  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-368 <HEI>  
A;Cross-references: UNIPROT:O9KT55; GB:AE0004186; GB:AE003852; NID:g9655507; PIDN:AAFF9420  
A;Experimental source: A82055; MUID:20406833; biotype El Tor  
C;Genetics:  
A;Gene: VC1050  
A;Map position: 1  
C;Superfamily: response regulator, Hnr type; response regulator homology

Query Match 37.1%; Score 53; DB 2; Length 368;  
Best Local Similarity 61.1%; Pred. No. 6.9; Matches 11; Conservative 3; Nismatches 4; Indels 0; Gaps 0;

Qy 9 ROYASALKHIIABIERGI 26  
Db 261 ENFSVELKDLABIERGI 278

**RESULT 4**  
T49463 hypothetical protein F8M21\_170 - *Arabidopsis thaliana*  
C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
C;Accession: T49463  
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Newes, H.W.; Rudd, S.; Le submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z24493  
A;Accession: T49463  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1227 <BEV>  
A;Cross-references: UNIPROT:Q9LXF4; EMBL:ALL353993; GSDB:GN00063; ATSP:F8M21\_170  
A;Experimental source: cultivar Columbia; BAC clone F8M21  
C;Genetics:  
A;Gene: ATSP:F8M21\_170  
A;Map position: 5

Query Match 36.4%; Score 52; DB 2; Length 1227;  
Best Local Similarity 33.3%; Pred. No. 33; Matches 9; Conservative 10; Nismatches 8; Indels 0; Gaps 0;

Qy 1 LALAYSSRQYASALKHIIABIERGIR 27  
Db 950 LVHGYSSEADISSSLR7LSAM1SKGMK 976

**RESULT 5**  
C64471 hypothetical protein H10920 homolog - *Methanococcus jannaschii*

C;Species: *Methanococcus jannaschii*  
C;Dte: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C;Accession: C64471  
C;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodde, A.; Iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A;Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.  
A;Reference number: A64300; MUID:96337999; PMID:668087  
A;Accession: C64471  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-280 <BLP>  
A;Cross-references: UNIPROT:Q58767; GB:U67577; NID:g2826400; PIDN:AAB99380\_1;  
C;Genetics:  
A;Map position: FOR1320764-1321606  
A;Start codon: GTG  
C;Superfamily: conserved hypothetical protein YBR002c

Query Match 35.7%; Score 51; DB 2; Length 280;  
Best Local Similarity 40.0%; Pred. No. 10; Matches 10; Conservative 7; Nismatches 8; Indels 0; Gaps 0;

Qy 1 LALAYSSRQYASALKHIIABIERG 25  
Db 174 IAIYGGQQEIIIDAVKIAEKVRKG 198

**RESULT 6**  
S37170 rggB protein - *Lactococcus lactis*  
C;Species: *Lactococcus lactis*  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S37170; S37169  
R;Seegers, J.F.M.J.; Franke, C.; Venema, G.; Bron, S.; Kiewiet, R.  
submitted to the EMBL Data Library, August 1993  
A;Description: The majority of lactococcal plasmids belong to one family.  
A;Reference number: S37169  
A;Accession: S37170  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-381 <SEE>  
A;Cross-references: UNIPROT:Q48703; EMBL:Z25477; NID:9401838; PIDN:CAA80966\_1; PID:94018:

Query Match 35.7%; Score 51; DB 2; Length 381;  
Best Local Similarity 33.3%; Pred. No. 14; Matches 9; Conservative 9; Nismatches 11; Indels 7; Gaps 0;

Qy 3 LAYYSSRQYASALKHIIABIERGIR 29  
Db 345 LSYVSSRQEAVSKRNVAKYLKKAIEQY 371

**RESULT 7**  
S40057 rggA protein - *Lactococcus lactis*  
C;Species: *Lactococcus lactis*  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S40057  
R;Frere, J.; Novel, M.; Novel, G.  
Mol. Microbiol. 10, 1113-1124, 1993  
A;Title: Molecular analysis of the *Lactococcus lactis* subspecies lactic plasmids  
A;Reference number: S40057; MUID:95020571; PMID:7934861  
A;Accession: S40057  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-388 <FRE>  
A;Cross-references: UNIPROT:Q48681; EMBL:X60454; NID:g414894; PIDN:CAA42983\_1; PID:9414861  
C;Superfamily: *Lactococcus lactis* replication protein repB

Query Match 35.7%; Score 51; DB 2; Length 388;

Qy	3 LAYSSRQASALKHAEIIEGIRQH 29	Score 8	Best Local Similarity 33.3%; Pred. No. 14; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;	Qy	11 YASLKHAEIIEGRI 26
Db	350 LSYVSSKQEAYSKRNYAKYLKKAEQY 376	RESULT 8	S37168 replication protein repB - Lactococcus lactis	Qy	11 YASLKHAEIIEGRI 26
			C;Species: Lactococcus lactis	Db	408 YAINLEHIGDIEKGII 423
			C;Accession: S37168; S44976		
			R;Seegers, J.; Franke, C.; Venema, G.; Bron, S.; Kiewiet, R.		
			submitted to the EMBL Data Library, August 1993		
			A;Description: The majority of lactococcal plasmids belong to one family		
			A;Reference number: S37168		
			A;Accession: S37168		
			A;Residues: 1-403 <SEBE>		
			A;Cross-references: UNIPROT:O48701; EMBL:Z25475		
			R;Seegers, J.; Franke, C.; Venema, G.; Bron, S.; Kiewiet, R.		
			submitted to the EMBL Data Library, February 1994		
			A;Description number: S44976		
			A;Accession: S44976		
			A;Status: preliminary		
			A;Molecule type: DNA		
			A;Residues: 1-103 <SEBE>		
			A;Cross-references: EMBL:Z25475; NID:9452758; PID:9452759		
			C;Genetics:		
			A;Gene: repB		
			C;Superfamily: Lactococcus lactis replication protein repB		
			Query Match Score 51; DB 2; Length 403;		
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	RESULT 9
			Query Match Score 51; DB 2; Length 403;		S37168 Agrobacterium tumefaciens (strain C58, Dupont)
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		AD2852 AgNPt/Pt cotransporter Atu2245 [Imported] - Agrobacterium tumefaciens
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	AD2852 Na+/Pt cotransporter Atu2245 [Imported] - Agrobacterium tumefaciens
			Query Match Score 51; DB 2; Length 403;		C;Accession: AD2852
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	R;Wood, D.W.; Setubal, J.C.; Kaufl, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
			Query Match Score 51; DB 2; Length 403;		erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutjavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S.
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		Science 294, 2317-2323, 2001
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, Ster, B.W.
			Query Match Score 51; DB 2; Length 403;		A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Reference number: AB2577; MUID:21608550; PMID:11743193
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	A;Accession: AD2852
			Query Match Score 51; DB 2; Length 403;		A;Status: preliminary
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Molecule type: DNA
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	A;Residues: 1-50 <KUR>
			Query Match Score 51; DB 2; Length 403;		A;Cross-references: UNIPROT:Q8UD82; GB:AE008688; PIDN:AAL43234_1; PID:g17740718; GSPDB:G
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Experimental source: strain C58 (Dupont)
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	C;Genetics:
			Query Match Score 51; DB 2; Length 403;		A;Gene: Auu2245
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Map position: circular chromosome
			Db	368 LSYVSSKQEAYSKRNYAKYLKKAEQY 394	RESULT 10
			Query Match Score 51; DB 2; Length 403;		B97629 hypothetical Protein AGR_C_4085 [imported] - Agrobacterium tumefaciens (strain C58, Cere)
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		C;Species: Agrobacterium tumefaciens
			Db	408 YAINLEHIGDIEKGII 423	C;Accession: B97629 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
			Query Match Score 51; DB 2; Length 403;		R;Goodner, B.; Hank, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Title: Genome sequence of the Plant Pathogen Agrobacterium tum
			Db	408 YAINLEHIGDIEKGII 423	A;Reference number: A97359; MUID:21608551; PMID:11743194
			Query Match Score 51; DB 2; Length 403;		A;Accession: B97629
			Best Local Similarity 33.3%; Pred. No. 15; Matches 9; Conservative 11; Mismatches 7; Indels 0; Gaps 0;		A;Status: preliminary
			Db	408 YAINLEHIGDIEKGII 423	A;Molecule type: DNA
			Query Match Score 51; DB 2; Length 403;		A;Residues: 1-555 <KUR>
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		A;Cross-references: UNIPROT:Q8UD82; GB:AE007869; PIDN:AAK87987_1; PID:g15157399; GSPDB:G
			Db	413 YAINLEHIGDIEKGII 428	C;Genetics:
			Query Match Score 51; DB 2; Length 403;		A;Gene: AGR_C_4085
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		A;Map position: circular chromosome
			Db	413 YAINLEHIGDIEKGII 428	RESULT 11
			Query Match Score 51; DB 2; Length 403;		H86257 protein F5011.2 [imported] - Arabidopsis thaliana
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		C;Species: Arabidopsis thaliana (mouse-ear cress)
			Db	413 YAINLEHIGDIEKGII 428	C;Accession: H86257
			Query Match Score 51; DB 2; Length 403;		R;Theologis, A.; Becker, J.R.; Palm, C.J.; Fedderspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		A;Authors: Hunter, J.L.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luces, J.S.; Maiti, R.; Marziali, A.;Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, J.;A;Authors: Salzberg, S.L.; Schwartz, J.R.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.;Ke, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.;A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis. A;Reference number: A86141; MUID:21016719; PMID:11130712
			Db	413 YAINLEHIGDIEKGII 428	A;Accession: H86257
			Query Match Score 51; DB 2; Length 403;		A;Status: preliminary
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		A;Molecule type: DNA
			Db	413 YAINLEHIGDIEKGII 428	A;Residues: 1-572 <STP>
			Query Match Score 51; DB 2; Length 403;		A;Cross-references: UNIPROT:Q9LN86; GB:AE005172; PIDN:AAF79628_1; GSPDB:GN
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		C;Genetics:
			Db	413 YAINLEHIGDIEKGII 428	A;Gene: F5011.2
			Query Match Score 51; DB 2; Length 403;		A;Map position: 1
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		Query Match Score 51; DB 2; Length 403;
			Db	413 YAINLEHIGDIEKGII 428	RESULT 12
			Query Match Score 51; DB 2; Length 403;		H30458 5 YYSSROYASALKHAEIIEGIRQH 29
			Best Local Similarity 33.3%; Pred. No. 21; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		C;Species: Sulfolobus solfatricus
			Db	413 YAINLEHIGDIEKGII 428	C;Species: Sulfolobus solfatricus

```

;Date: 24-May-2001 # sequence revision 24-May-2001 #text_change 09-Jul-2004
;Accession: R90458
;Orgel: Q ; Singh, R.K.; Conflalonieri, F.; Zivanovic, Y.; Allard, G.; Awatez, M.J.; Chan-Jong, I.; Jeffries, A.C.; Kozer, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, E.; Rettar, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
;Submitted to GenBank: April 2001
;Description: Sulfolobus solfataricus complete genome.
;Reference number: 99139
;Accession: R90458
;Status: preliminary
;Molecule type: DNA
;Cross-references: UNIPROT:Q97V20 ; GB:AE006641; PIDN:ARK42927.1; GSPDB:Q
;Genetics:
;Gene: SSO2816
Query Match      35.3%; Score 50.5%; DB 2; Length 316;
Best Local Similarity 40.4%; Pred. No. 14;
Matches 13; Conservative 7; Mismatches 5; Gaps 1;
Y :|||: :|||: :|||: :|||: :|||: 1
b 164 TAVATSSGGTEVARGAYDVGHLKEITKKAK 195
1 LALAYS-----SROYASALKHAIEELRGIR 27

```

RESULT 13  
 69822  
 ABC transporter (ATP-binding protein) homolog yhCH - *Bacillus subtilis*  
 ;Species: *Bacillus subtilis*  
 ;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
 ;Accession: D69822  
 ;Kunst, F.; Ogasawara, N.; Moszer, I.; Alberini, A.M.; Azevedo, V.; Berter, J.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc, S.D.; Ehrlich, P.R.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Galler, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hull, M.F.; Jardinois, P.; Koningschein, G.; Krogh, S.; Kurita, K.; Lapidus, A.; Lardinois, M.; Levine, A.; Liu, H.; Masuda, S.; Mause, T.M.; Portetelle, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Scanlon, Rieder, M.; Rivolt, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaike, Y.; Sato, T.; Serov, A.; Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekowska, J.; Seto, T.; Tegoni, A.; Tosato, V.; Uchiyama, K.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; Yamane, K.; Yamamoto, H.; Yoshikawa, E.; Yoshida, K.; Winters, P.; Wibat, A.; Yamamoto, K.; Yata, K.; Yoshida, H.; Zimstein, H.F.; Danchin, A.; Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. Reference number: A69580; PMID: 98044033; MUID: 934377

Scoring table:	BLOSUM62	Alignments		
Gapop:	10.0	Gapext 0.5		
Searched:	1825181 seqs, 575374646 residues	SEQUENCE FROM N.A.		
Total number of hits satisfying chosen parameters:	1825181	SEQUENCE=Jung;		
Minimum DB seq length:	0	RC		
Maximum DB seq length:	2000000000	PubMed=14703039;		
Post-processing:	Minimum Match 0%	RA		
	Maximum Match 100%	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,		
Database :	UniProt 02: 1: uniprot_sprot: 2: uniprot_trmbi:*	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wasatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Taniai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii K., Hara T., Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Morikawa S., Momiyama H., Sato N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki K., Watanabe T., Sugiyama A., Itakura S., Fukuzumi Y., Yamazaki M., Watanabe K., Kumada A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Morishima-Sugano J., Satoh H., Watanabe M., Komatsu T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S., Oritani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai T., Takahashi Y., Nakagawa K., Okunura K., Nagase T., Nomura N., Kakuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Israi T., Sugano S., "Complete sequencing and characterization of 21,243 full-length human cDNAs", Nat. Genet. 36:40-45 (2004).		
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	RN		
SUMMARIES	%	RESULT 1		
Result No.	Score	Query Match Length DB ID	Description	
-	-	-	-	-
1	143	100.0	638 2 Q96NE6	Q96NE6 homo sapien
2	143	100.0	665 2 Q81VP2	Q81VP2 homo sapien
3	143	100.0	665 2 Q8N4P2	Q8N4P2 homo sapien
4	143	100.0	678 2 Q86WT1	Q86WT1 homo sapien
5	94.4	664 2 Q8C0X1	Q8C0X1 mus musculus	
6	135	94.4	664 2 Q99J3B	Q99J3B m riken cdn
7	135	94.4	664 2 Q8CY00	Q8CY00 mus musculus
8	118	82.5	651 2 Q6INU8	Q6INU8 xenopus laevis
9	118	82.5	651 2 AAH72174	AAH72174 xenopus laevis
10	81	56.6	634 2 Q1PRA4	Q1PRA4 aphelinus
11	78	54.5	652 2 Q8VK41	Q8VK41 drosophila
12	76	53.1	597 2 Q817G3	Q817G3 caenorhabditis elegans
13	76	53.1	656 2 Q817G4	Q817G4 caenorhabditis elegans
14	55	38.5	516 2 Q9WB1	Q9WB1 agrobacteri
15	54	37.8	131 2 Q57408	Q57408 aquifex aeolicus
16	54	37.8	209 2 Q8PBPS	Q8PBPS xanthomonas campestris
17	54	37.8	209 2 Q8PNA0	Q8PNA0 xanthomonas campestris
18	54	37.8	663 2 Q8MN61	Q8MN61 dictyostelia
19	54	37.8	1279 2 Q7QV8	Q7QV8 giardia lamblia
20	53	37.1	368 2 Q7MIU5	Q7MIU5 vibrio vulnificus
21	53	37.1	368 2 Q8BB27	Q8BB27 vibrio vulnificus
22	53	37.1	368 2 Q8KT55	Q8KT55 vibrio cholerae
23	52	36.4	247 2 Q8BSM9	Q8BSM9 methanospaerulum
24	52	36.4	365 2 Q7NOT3	Q7NOT3 photorhabdus luminescens
25	52	36.4	495 1 T8SB_SULAC	T8SB_SULAC
26	52	36.4	595 1 Q8PZ00	Q8PZ00 drosophilidae
27	52	36.4	615 2 Q93Y16	Q93Y16 drosophilidae
28	52	36.4	615 2 AAFA6175	AAFA6175 drosophilidae
29	52	36.4	1227 2 Q9LXF4	Q9LXF4 arabidopsis thaliana
30	51.5	36.0	101 2 Q9DA67	Q9DA67 mus musculus
31	51.5	36.0	226 2 Q8C1D2	Q8C1D2 mus musculus

Repeat; TPR repeat.

SQ	SEQUENCE	638 AA;	73146 MW;	16AB964E71ACF893 CRC64;				
	Query Match	100.0%;	Score 143;	DB 2;	Length 638;			
	Best Local Similarity	100.0%;	Pred. No. 6.1e-12;					
	Matches	29;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1	LALAYSSROYASALKHIAETTERGIRQH	29					
Db	167	LALAYSSROYASALKHIAETTERGIRQH	195					
<b>RESULT 2</b>								
Q8IVVP2		PRELIMINARY;		PRT;	665 AA.			
ID	Q8N4P2;			PRELIMINARY;		PRT;	665 AA.	
AC	Q8N4P2;							
DT	01-OCT-2002	(TREMBLrel.	22;	Created)				
DT	01-OCT-2002	(TREMBLrel.	22;	Last sequence update)				
DT	01-MAR-2004	(TREMBLrel.	26;	Last annotation update)				
DE	FLJ30990 protein.							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
NCBI_TAXID=9606;								
OX								
RN	[1]							
<b>SEQUENCE FROM N.A.</b>								
RP	PTTUE-BRAIN;							
RC	MEDLINE-22388257;							
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,							
RA	Klausner R.D., Collins F.S., Wagner L., Shevmen C.M., Schuler G.D.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loqueland N.A., Peters G.J., Malek J.A., Gunarathne P.H.,							
RA	McEvany P.J., McKernan K.J., Garcia A.M., Hale S., Worley K.C.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,							
RA	Fahay J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C							



AC Q99J38;	SEQUENCE FROM N.A.
DT 01-JUN-2001 (TRIMBLrel. 17; Created)	RP SEQUENCE FROM N.A.
DT 01-JUN-2001 (TRIMBLrel. 17; Last sequence update)	RP STRAIN=C57BL/6J; TISSUE=Parthenogenote;
DT 01-OCT-2004 (TRIMBLrel. 28; Last annotation update)	RC MEDLINE=20303913; PubMed=1107861;
DB RIKEN CDNA 4930506Li3 (Mus musculus 9.5 days embryo Parthenogenote	RA Striata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
DB RIKEN full-length enriched library, Clone:RJ03032P09	RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
DB product:hydrochemical tetrapeptide repeat (IPR) structure	RA Sumi N., Ishii Y., Nakamura S., Hizama M., Kishine T., Hazada A.,
DB containing protein, full insert sequence).	RA Yamamoto R., Matsumoto S., Sakaguchi S., Ikegami T., Kashiwagi K.,
DB Name=4930506Li3Rik	RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
GN Mus musculus (Mouse).	RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
OS Mammalia: Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi;	RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
OC Mammalia: Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.	RT "RIKEN integrated sequence analysis (RISA) system-384" format
NCBI_TaxID=10090;	RT sequencing pipeline with 384 multicapillary sequencer.;"
RN	RT Genome Res. 10:1757-1771 (2000).
RP	RN [1]
SEQUENCE FROM N.A.	RN [8]
RC STRAIN=NMR1; TISSUE=Mammary tumor. WAP-Tag model. 5 months old;	RN SEQUENCE FROM N.A.
RX MEDLINE=2238257;	RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derege J.G., Schuller G.D.,	RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schaefer C.P., Bhat N.K.,	RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Altschul S.F., Zeerberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,	RA Hayashida K., Hayatsu N., Hiramoto K., Hiraduka T., Hirozane T.,
RA Hopkins R.P., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,	RA Hori F., Imotani K., Ishii Y., Itch M., Kageya I., Kasukawa T.,
RA Diatchenko L., Mansurina K., Farmer A.A., Rubin G.M., Hong L.,	RA Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Roya S.,
RA Stapleton M., Soresi M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	RA Kuribara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	RA Nishi K., Nomura K., Numasaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Raha S.S., Loqueline N.A., Peters G.J., Abramson R.D., Mullahay S.J.,	RA Saito R., Saithoh H., Sakai C., Sakai K., Sakazume N., Sano H., Tagami M.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tomaru A., Tomaru A., Toya T., Yasunishi A., Muranatsu M., Hayashizaki Y.,
RA Villalon D.K., Muzny D.M., Soedergren E.J., Lu X., Gibbs R.A.,	RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RA Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,	DR EMBL; BC004773; AAH04773_1;
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	DR EMBL; BC045196; BAC3220_1;
RA Blodgesley R.W., Touchman J.W., Greenblatt J., Myers R.M., Butterfield Y.S.,	DR MGD; MGI: 1226052; 4330505113Rik.
RA Rodriguez A.C., Grinwood J., Schmitz J., Smailus D.B., Schnarch A., Schein J.E.,	DR InterPro; IPR008940; Prenyl_trans.
RA Krzywinski M.I., Skalska U., Smailus D.B., Schnarch A., Schein J.E.,	DR InterPro; IPR004440; TPR.
RA Jones S.J., Marra M.A.;	DR Pfam; PF00515; TPR; 4.
RT "Generation and initial analysis of more than 15,000 full-length human	DR SMART; SM00028; TPR; 4.
RT and mouse cDNA sequences";	DR PROSITE; PS05293; TPR_repeat; TPR repeat.
PRO. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR NW Hypothetical protein; Repeat; TPR repeat.
RN [2]	DR NW Score 135; DB 2; Length 664;
SEQUENCE FROM N.A.	SEQ SEQUENCE 664 AA: 76237 MW: F9C3C783231CBA2 CRC64;
RC STRAIN=NMR1; TISSUE=Mammary tumor. WAP-Tag model. 5 months old;	Qy 1 Lalayissroyasalrhiaettiergrqh 29
RA Strauberg R.; Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.	DD 193 Lalayissroyapalkhiadiergirqh 221
RN [3]	RESULT 7
SEQUENCE FROM N.A.	Q9CY00 PRELIMINARY; PRT; 664 AA.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;	AC Q9CY00; PRT; 664 AA.
RX MEDLINE=21085660; PubMed=11217851;	DT 01-JUN-2001 (TRIMBLrel. 17, Created)
RA RIKEN FANTOM Consortium;	DT 01-JUN-2001 (TRIMBLrel. 17, Last sequence update)
RT "Functional annotation of a full-length mouse cDNA collection.";	DT 01-OCT-2003 (TRIMBLrel. 25, Last annotation update)
RL Nature 409:685-690 (2001).	DE Mus musculus 13 days embryo liver cDNA, RIKEN full-length enriched
RN [5]	DE Library, clone:510042P03 product:4930506Li3Rik PROTEIN homolog.
SEQUENCE FROM N.A.	GN Name=2510042P03Rik;
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;	OS Mus musculus (Mouse).
RA Carninci P., Hayashizaki Y.;	OC Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi;
RT "High-efficiency full-length cDNA cloning.";	OC NCBI_TaxID=10090;
RL Meth. Enzymol. 303:19-44 (1999).	RN [1]
RN [4]	RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;	RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=99279253; PubMed=10349636;	RC MEDLINE=99279251; PubMed=110349636;
RA Carninci P., Hayashizaki Y.;	RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";	RT "High-efficiency full-length cDNA cloning.";
RT Meth. Enzymol. 303:19-44 (1999).	RT Meth. Enzymol. 303:19-44 (1999).
RN [6]	RT [2]
SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;	RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=20169937; PubMed=11042559;	RC MEDLINE=99279251; PubMed=110349636;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,	RA Carninci P., Hayashizaki Y.;
RA Kondo H., Okazaki Y., Muramatsu M., Hayashizaki Y.;	RA Kondo H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to	RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630 (2000).	RT Genome Res. 10:1617-1630 (2000).

RX MEDLINE=21085660; PubMed=11217851;	DT 05-JUL-2004 (TREMBLrel. 27, Created)
RT "Functional annotation of a full-length mouse cDNA collection.";	DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
RN [3]	DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
RP SEQUENCE FROM N.A.	DB MGCB0272 protein.
RC STRAIN=C57BL/6J; TISSUE=Liver;	GN Name=MGE80272;
RA THE FANTOM Consortium,	OS Xeropus laevis (African clawed frog)
the RIKEN Genome Exploration Research Group Phase I & II Team;	OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
"Analysis of the mouse transcriptome based on functional annotation of	OC Xenopodinae; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
60,770 full-length cDNAs.";	OC Xenopus.
RT RN NCBI_TaxID=3455;	OX
RN [4]	RN
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;	RC TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159;	RX MEDLINE=22389257; PubMed=12477932;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,	RA Klausner R.D., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;	RA Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to	RA Mung S.T., Moore T., Max G.M., Hsieh F.,
RT prepare full-length cDNA libraries for rapid discovery of new genes.";	RA Diachenko L., Marusina K., Farmer A.A., Rubin M.F., Horsch L.,
RL Genome Res. 10:1617-1630 (2000).	RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RN [5]	RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RP SEQUENCE FROM N.A.	RA Raha S.S., Loqueilano N.A., Peters G.J., Carninci P.D., Mullany S.J.,
RC STRAIN=C57BL/6J; TISSUE=Liver;	RA Boesk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RX MEDLINE=20530913; PubMed=11076861;	RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,	RA Villalon D.M., Muzyz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Konno H., Akiyama J., Nishii K., Kitesuna T., Tashiro H., Itoh M.,	RA Fahey J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
RA Sumi N., Ishii Y., Nakamura S., Razama M., Nishine T., Harada A.,	RA Whiteley M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,	RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Fujikawa S., Inoue K., Togawa Y., Izawa M., Obara E., Watanuki M.,	RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,	RA Krzywinski M.I., Skalska U., Smailis D.E., Schein J.E.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;	RA Jones S.J., Marr A.M., RT "Generation and initial analysis of more than 15,000 full-length human
RP sequencing pipeline with 384 multicapillary sequencer.";	RT and mouse cDNA sequences";
RL Genome Res. 10:1757-1771 (2000).	RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [6]	RN [2]
RP SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;	RC TISSUE=Ovary;
RA Adachi J., Aizawa K., Akaiwa S., Akimura T., Arai A., Aono H.,	RX MEDLINE=22341132; PubMed=12454917;
RA Arakawa T., Bono H., Carninci P., Fukuda S., Furukoshi Y., Furuno M.,	RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Hanagata T., Hara A., Hayatsu N., Hiraoaka T., Hiraoka T., Hori F.,	RA Richardson P., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,	RA RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RA Kawai J., Koijima Y., Konno H., Kouda M., Koya S., Kurihara C.,	RA RT initiative." Dev. Dyn. 225:384-391 (2002).
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numasaki R., Onno M.,	RL Dev. Dyn. 225:384-391 (2002).
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,	RN [3]
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,	RP SEQUENCE FROM N.A.
RA Tejima Y., Suzuki H., Tagami M., Tagashira F., Tanaka T., Yoshino M.,	RC TISSUE=Stratified squamous epithelial tissue;
RA Muramatsu M., Hayashizaki Y.;	RX MEDLINE=22341132; PubMed=12454917;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.	RA Klein S., Strausberg R.,
DR EMBL; AK011097; BAB2394.; -	RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR InterPro; IPR001440; TPR; 4.	DR EMBL; BC072174; AAH72174.1; -.
DR InterPro; IPR00841; TPR-like.	DR InterPro; IPR001440; TPR; 2.
DR SMART; SM00028; TPR; 4.	DR InterPro; IPR00841; TPR-like.
DR PROSITE; PSS0293; TPR_REGION; 1.	DR SMART; SM00028; TPR; 2.
KW Repeat; TPR repeat.	DR PROSITE; PS50005; TPR; 1.
SQ SEQUENCE 664 AA; 76095 MW; 8BCB2D1FA7099F3 CRC64;	DR PROSITE; PS50093; TPR_REGION; 2.
Query Match 94.4%; Score 135; DB 2; Length 664;	DR Repeat; TPR repeat.
Best Local Similarity 93.1%; Pred. No. 9.2e-11;	DR PROSITE; PS50093; TPR_REGION; 2.
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	SQ SEQUENCE 651 AA; 75558 MW; AF0608F6FF7986590 CRC64;
Qy 1 LALAYSSROYASALKHIAELIERGIROH 29	Query Match 92.5%; Score 118; DB 2; Length 651;
Db 193 LALAYSSROYASALKHIAELIERGIROH 221	Best Local Similarity 75.9%; Pred. No. 2.6e-08;
RESULT 8	Matches 22; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QINUS PRELIMINARY;	Qy 1 LALAYSSROYASALKHIAELIERGIROH 29
ID Q6INU8 ID AAH72174	Db 180 LALCYSSROYASALKHADIERGIROH 208
AC Q6INU8 AC AAH72174	DT 01-TUN-2004 (TREMBLrel. 27, Created)
	DT 01-TUN-2004 (TREMBLrel. 27, Last sequence update)







Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw mode!

Run On: November 10, 2004, 13:44:14 ; Search time 13.9094 Seconds

(without alignments)

181.178 Million cell updates/sec

Title: US-10-092-750-46

Perfect score: 210

Sequence: 1 AAMLDREGTECDIWINEMSLHKTIVQDVYGTPHPPHS 38

Scoring table: BLOSUM62

Gapopen 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters:

478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:  
 1: /cgn2\_6/ptodata/1/iaas/5A\_COMB\_pep:  
 2: /cgn2\_6/ptodata/1/iaas/5B\_COMB\_pep:  
 3: /cgn2\_6/ptodata/1/iaas/6A\_COMB\_pep:  
 4: /cgn2\_6/ptodata/1/iaas/6B\_COMB\_pep:  
 5: /cgn2\_6/ptodata/1/iaas/PCUTS\_COMB\_pep:  
 6: /cgn2\_6/ptodata/1/iaas/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Description
			%
1	57.5	27.4	US-09-270-767-46552
2	56	26.7	Sequence 46552, A
3	52	24.8	Sequence 4753, Ap
4	52	24.8	Sequence 3745, Ap
5	52	24.8	Sequence 1037, Ap
6	51.5	24.5	Sequence 55, Ap
7	51	24.3	Sequence 261, Ap
8	51	24.3	Sequence 298, Ap
9	51	24.3	Sequence 566, Ap
10	51	24.3	Sequence 567, Ap
11	51	24.3	Sequence 567, Ap
12	51	24.3	Sequence 567, Ap
13	51	24.3	Sequence 567, Ap
14	51	24.3	Sequence 567, Ap
15	51	24.3	Sequence 567, Ap
16	51	24.3	Sequence 567, Ap
17	50.5	24.0	Sequence 263, Ap
18	50.5	24.0	Sequence 303, Ap
19	50.5	24.0	Sequence 1431, Ap
20	50	23.8	Sequence 178, Ap
21	50	23.8	Sequence 649, Ap
22	49	23.3	Sequence 138, Ap
23	49	23.3	Sequence 138, Ap
24	49	23.3	Sequence 152, Ap
25	49	23.3	Sequence 197, Ap
26	49	23.3	Sequence 517, Ap
27	49	23.3	Sequence 2037, Ap

## DESCRIPTION

Result No.	Score	Query Match	Description
			%
1	57.5	27.4	US-09-270-767-46552
2	56	26.7	Sequence 46552, A
3	52	24.8	Sequence 4753, Ap
4	52	24.8	Sequence 3745, Ap
5	52	24.8	Sequence 1037, Ap
6	51.5	24.5	Sequence 55, Ap
7	51	24.3	Sequence 261, Ap
8	51	24.3	Sequence 298, Ap
9	51	24.3	Sequence 566, Ap
10	51	24.3	Sequence 567, Ap
11	51	24.3	Sequence 567, Ap
12	51	24.3	Sequence 567, Ap
13	51	24.3	Sequence 567, Ap
14	51	24.3	Sequence 567, Ap
15	51	24.3	Sequence 567, Ap
16	51	24.3	Sequence 567, Ap
17	50.5	24.0	Sequence 263, Ap
18	50.5	24.0	Sequence 303, Ap
19	50.5	24.0	Sequence 1431, Ap
20	50	23.8	Sequence 178, Ap
21	50	23.8	Sequence 649, Ap
22	49	23.3	Sequence 138, Ap
23	49	23.3	Sequence 138, Ap
24	49	23.3	Sequence 152, Ap
25	49	23.3	Sequence 197, Ap
26	49	23.3	Sequence 517, Ap
27	49	23.3	Sequence 2037, Ap

Result No.	Score	Query Match	Description
			%
1	28	4.9	Sequence 2414, Ap
2	29	4.9	Sequence 2414, Ap
3	30	4.9	Sequence 2414, Ap
4	31	4.9	Sequence 2414, Ap
5	32	4.9	Sequence 2441, Ap
6	33	4.9	Sequence 2441, Ap
7	34	4.9	Sequence 2441, Ap
8	35	4.9	Sequence 2442, Ap
9	36	4.9	Sequence 2442, Ap
10	37	4.8	Sequence 2425, Ap
11	38	4.8	Sequence 520, Ap
12	39	4.8	Sequence 565, Ap
13	40	4.8	Sequence 1288, Ap
14	41	4.8	Sequence 1288, Ap
15	42	4.8	Sequence 1911, Ap
16	43	4.8	Sequence 1939, Ap
17	44	4.8	Sequence 1971, Ap
18	45	4.8	Sequence 1999, Ap

## ALIGNMENTS

## RESULTS

RESULT 1  
US-09-270-767-46552

; Sequence 46552, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-034

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 6351-7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 46552

; LENGTH: 1124

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

; US-09-270-767-46552

; Query Match

; Best Local Similarity

;保守性

; Matches

; Score

; DB

; Indels

; Gaps

; 2;

; Qy

; 5 LDPRRGTCDCDILWIMENSLHKTIVQDVYGT

; PHPPHS 38

; Db

; 569 IDRQSSRL-CWLS-----DTYVTPHPHPS 591

; RESULT 2

US-09-543-681A-4753

; Sequence 4753, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709-1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO: 4753

; LENGTH: 334

; TYPE: PRT

; ORGANISM: Proteus mirabilis

; US-09-543-681A-4753

; Query Match

```

Db      10 DDGGRELLAWIIRREMRRHHPVHEDYGAFF 38
Best Local Similarity 48.0%; Pred. No. 8.2;          Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
Query 5 LDREGTECDLWINEMSLLHKIVQDV 29
Db      52 LDYRKPKSDQWINGTELLPIAHDV 76

RESULT 3
US-09-540-236-3745
; Sequence 3745, Application US/09540236
; GENERAL INFORMATION:
;   Patent No. 6673910
;   APPLICANT: Gary L. Breton et al.
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
;   TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;   FILE REFERENCE: 2709.2005-001
;   CURRENT APPLICATION NUMBER: US/09/540.236
;   CURRENT FILING DATE: 2000-04-04
;   NUMBER OF SEQ ID NOS: 3840
;   SEQ ID NO: 3745
;   LENGTH: 219
;   TYPE: PRT
;   ORGANISM: M.catarrhalis
;   SEQ ID NO: 55

Query Match 24.8%; Score 52; DB 4; Length 219;
Best Local Similarity 27.0%; Pred. No. 18; Mismatches 5; Indels 13; Gaps 3;
Matches 17; Conservative 17;

Query 1 AAMILDRRETECDLW--INEMSLHKIV-----VQDVYGT 33
Db      49 ALMADAKG-ECDYWLRINAHVSHMHLDDITCMKLKUNHLLGVVLPKCQNKHQSVEVYHHI 107
Query 34 HPP 36
Db      108 HKP 110

RESULT 4
US-08-765-907A-10
; Sequence 10, Application US/08765907A
; GENERAL INFORMATION:
;   Patent No. 6352339
;   APPLICANT: BLANC, Veronique
;   APPLICANT: THIBAUT, Denis
;   APPLICANT: Bamas-JACQUES, Nathalie
;   APPLICANT: BLANCHE, Francis
;   APPLICANT: DUTRUC-ROSSET, Gilles
;   APPLICANT: BAIZET, Joel
;   APPLICANT: BARRIERE, Jean-Claude
;   APPLICANT: DEBUSSCHE, Laurent
;   APPLICANT: FAMBACH, Alain
;   APPLICANT: PARIS, Jean-Marc
;   TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
;   TITLE OF INVENTION: Mutasynthesis
;   FILE REFERENCE: Streptogramin genes
;   CURRENT APPLICATION NUMBER: US/08/765, 907A
;   CURRENT FILING DATE: 1997-03-20
;   NUMBER OF SEQ ID NOS: 17
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO: 10
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: Streptomyces pristinaespiralis
;   SEQ ID NO: 10

Query Match 24.8%; Score 52; DB 3; Length 399;
Best Local Similarity 37.9%; Pred. No. 36; Mismatches 3; Indels 15; Gaps 0;
Matches 11; Conservative 3;

RESULT 5
US-09-340-620A-55
; Sequence 55, Application US/09340620A
; GENERAL INFORMATION:
;   Patent No. 646233
;   APPLICANT: Berlin, John
;   TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
;   FILE REFERENCE: 07334-124001
;   CURRENT APPLICATION NUMBER: US/09/340.620A
;   CURRENT FILING DATE: 1999-06-28
;   PRIOR APPLICATION NUMBER: US 09/245, 281
;   PRIOR FILING DATE: 1999-02-05
;   PRIOR APPLICATION NUMBER: US 09/207, 359
;   PRIOR FILING DATE: 1998-12-08
;   PRIOR APPLICATION NUMBER: US 09/099, 041
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: US 09/019, 942
;   PRIOR FILING DATE: 1998-02-06
;   NUMBER OF SEQ ID NOS: 71
;   SOFTWARE: Fastseq for Windows Version 4.0
;   SEQ ID NO: 55

Query Match 24.8%; Score 52; DB 4; Length 1037;
Best Local Similarity 36.4%; Pred. No. 1.1e+02;
Matches 12; Conservative 6; Mismatches 7; Indels 8; Gaps 1;
Query 1 AAMILDRRETECDLWINEMSLHKIVQDVYGT 33
Db      698 ALMPIQEPCTQCELSQN-----LQNLGYTP 722

RESULT 6
US-09-377-285B-44
; Sequence 44, Application US/09377285B
; Patent No. 672075
; GENERAL INFORMATION:
;   APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;   APPLICANT: WOLEY, Paul
;   APPLICANT: TU, Jian
;   APPLICANT: XIAO, Bo
;   APPLICANT: LEAHY, Daniel
;   APPLICANT: BEEREN, Anthony
;   APPLICANT: LANAHAN, Anthony
;   TITLE OF INVENTION: NUCLEAR ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
;   CURRENT APPLICATION NUMBER: THU150-4
;   CURRENT FILING DATE: 1999-08-18
;   FILE REFERENCE: THU150-4
;   CURRENT APPLICATION NUMBER: US/09/377, 285B
;   CURRENT FILING DATE: 1999-08-18
;   FILE REFERENCE: THU150-4
;   PRIOR APPLICATION NUMBER: US 60/138, 426
;   PRIOR FILING DATE: 1999-06-10
;   PRIOR APPLICATION NUMBER: US 60/138, 493
;   PRIOR FILING DATE: 1999-06-10
;   PRIOR APPLICATION NUMBER: US 60/138, 494
;   PRIOR FILING DATE: 1999-06-10
;   PRIOR APPLICATION NUMBER: US 60/097, 334
;   PRIOR FILING DATE: 1998-08-18
;   NUMBER OF SEQ ID NOS: 72
;   SOFTWARE: PatentIn version 3.0
;   SEQ ID NO: 44
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
;   SEQ ID NO: 44

Query Match 24.5%; Score 51.5; DB 4; Length 261;
Best Local Similarity 32.4%; Pred. No. 26; Mismatches 9; Indels 1; Gaps 1;
Matches 11; Conservative 3;

Query Match 24.8%; Score 52; DB 3; Length 399;
Best Local Similarity 37.9%; Pred. No. 36; Mismatches 3; Indels 15; Gaps 0;
Matches 11; Conservative 3;

```

Query 5 LDRRGTCEDLWIN-EMSILHKIVQDVYGTPHPPH 37  
 Database 103 LERRGEEREMLVDEEYEYQDRLRDMEAHPHPPH 136

RESULT 7  
 US-09-248-796A-14868  
 Sequence 14868, Application US/09248796A  
 Patent No. 674137  
 GENERAL INFORMATION:  
 APPLICANT: Keith Weinstock et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 FILE REFERENCE: 101196.132  
 CURRENT FILING DATE: 1999-02-12  
 PRIORITY NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIORITY NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 14868  
 LENGTH: 298  
 TYPE: PRT  
 ORGANISM: Candida albicans  
 US-09-248-796A-14868

Query Match 24.3%; Score 51; DB 4; Length 298;  
 Best Local Similarity 28.6%; Pred. No. 36;  
 Matches 12; Conservative 5; Mistmatches 11; Indels 14; Gaps 2;

Qy 9 GTCDCDL-----W--INEMSLHKIVQDVYGTPHPP 36  
 Db 74 GTESSIONQHHAHHSNLFWKLNFNQSQQLHKVISONRNSPHEP 115

RESULT 8  
 US-09-028-934-31  
 Sequence 31, Application US/09028934  
 Patent No. 611670  
 GENERAL INFORMATION:  
 APPLICANT: Ligon, James M.  
 APPLICANT: Hill, Dwight S.  
 APPLICANT: Lam, Steven T.  
 APPLICANT: Hammer, Phillip E.  
 APPLICANT: van Pee, Karl Heinz  
 APPLICANT: Kirner, Sabine  
 APPLICANT: Young, Thomas R.  
 TITLE OF INVENTION: Pyrrolinitrin Biosynthesis Genes and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6117670artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/028,934  
 FILING DATE:

Filing Date: 09-OCT-1996  
 Prior Application Data:  
 Application Number: US 08/258,261

FILING DATE: 08-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8639  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 566 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-028-934-31

RESULT 9  
 US-08-258-261B-4  
 Sequence 4, Application US/08258261B  
 Patent No. 5639949  
 GENERAL INFORMATION:  
 APPLICANT: Schupp, Thomas  
 APPLICANT: Ligon, James M.  
 APPLICANT: Beck, James Joseph  
 APPLICANT: Hill, Dwight Steven  
 APPLICANT: Ryals, John Andrew  
 APPLICANT: Gaffney, Thomas Deane  
 APPLICANT: Lam, Stephen Ting  
 APPLICANT: Hammer, Phillip E.  
 APPLICANT: Ukenes, Scott Joseph  
 TITLE OF INVENTION: Genes for the synthesis of  
 antipathogenic substances  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ciba-Geigy Corporation  
 STREET: 7 Skyline Drive  
 CITY: Hawthorne  
 STATE: NJ  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/258,261B  
 FILING DATE: 08-JUN-1994  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/457,205  
 FILING DATE: 01-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8889  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 567 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-258-261B-4

Query Match Similarity 24.3%; Score 51; DB 1; Length 567;  
Best Local Similarity 45.5%; Pred. No. 74;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 16 INEMSLIHKIVQDVYGTPHPPH 37  
Db 58 IPETLANRRIADRYGIPELDH 79

RESULT 10

Sequence 4, Application US/08456837  
Patent No. 563374

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Uknas, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, Dwight Steven  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Uknas, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NJ  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/457,342  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/457,205  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/456,837  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: US/08/258,261  
FILING DATE: 08-Jun-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 567 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-456-837-4

RESULT 11

Sequence 4, Application US/08457342  
Patent No. 5662898

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, Dwight Steven  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Uknas, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, Dwight Steven  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew

TITLE OF INVENTION: Genes for the synthesis of  
antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/457,342  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/457,205  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/456,837  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: US/08/258,261  
FILING DATE: 08-Jun-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 567 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-457-342-4

RESULT 12

Sequence 4, Application US/08457646A  
Patent No. 5675560

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, Dwight Steven  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew

Query Match Similarity 24.3%; Score 51; DB 1; Length 567;  
Best Local Similarity 45.5%; Pred. No. 74;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 16 INEMSLIHKIVQDVYGTPHPPH 37  
Db 58 IPETLANRRIADRYGIPELDH 79

Query Match Similarity 24.3%; Score 51; DB 1; Length 567;  
Best Local Similarity 45.5%; Pred. No. 74;  
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 16 INEMSLIHKIVQDVYGTPHPPH 37  
Db 58 IPETLANRRIADRYGIPELDH 79



CLASSIFICATION: 600  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/457,205  
 FILING DATE: 01-JUN-1995  
 APPLICATION NUMBER: 08/258,261  
 FILING DATE: 08-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIPS  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-6689  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 567 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-457-335A-4

Query Match 24.3%; Score 51; DB 2; Length 567;  
 Best Local Similarity 45.5%; Pred. No. 74;  
 Matches 10; Conservative 4; Mismatches 8; Indels 0;  
 Gaps 0;

Qy 16 INEMSLIHKIVDGYGTPHPH 37  
 Db 58 IPETSLNRIIADRYGIPELDH 79

RESULT 15  
 US-08-729-214-4  
 Sequence 4, Application US/08729214  
 Patent No. 581502  
 GENERAL INFORMATION:  
 APPLICANT: Ligon, James M.  
 APPLICANT: Hill, Dwight Steven  
 APPLICANT: Ryals, John Andrew  
 APPLICANT: Hammer, Phillip E.  
 APPLICANT: van Pee, Karl Heinz  
 APPLICANT: Kirner, Sabine  
 TITLE OF INVENTION: Genes for the synthesis of  
 TITLE OF INVENTION: antipathogenic substances  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ciba-Geigy Corporation  
 STREET: 520 White Plains Road  
 CITY: Tarrytown  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10591  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentnet Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/729,214  
 FILING DATE: TBA  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIPS  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 567 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

DN protein - protein search, using sw model  
 run on: November 11, 2004, 01:28:30 ; Search time 43.9509 Seconds  
 (without alignments)  
 305,399 Million cell updates/sec

Query: US-10-092-750-46

Perfect score: 210 Sequence: 1 AAMLDRGTECDLWINEMSLHKKIVQDVYGTPHPHHS 38

Scoring table: BLOSUM62 Gapext 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353223886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries

Published Applications AA:  
 1: /cgn2\_6/ptodata/1/pubpa/us07\_PUBCOMB.pep/\*  
 2: /cgn2\_6/ptodata/1/pubpa/pct\_NW\_PUB.pep/\*  
 3: /cgn2\_6/ptodata/1/pubpa/us06\_\_NEW\_PUB.pep/\*  
 4: /cgn2\_6/ptodata/1/pubpa/us06\_PUBCOMB.pep/\*  
 5: /cgn2\_6/ptodata/1/pubpa/us07\_NEW\_PUB.pep/\*  
 6: /cgn2\_6/ptodata/1/pubpa/pcetus\_PUBCOMB.pep/\*  
 7: /cgn2\_6/ptodata/1/pubpa/us08\_NEW\_PUB.pep/\*  
 8: /cgn2\_6/ptodata/1/pubpa/PUBCOMB.pep/\*  
 9: /cgn2\_6/ptodata/1/pubga/us09A\_PUBCOMB.pep/\*  
 10: /cgn2\_6/prodata/1/pubpa/us09B\_PUBCOMB.pep/\*  
 11: /cgn2\_6/prodata/1/pubpa/us09C\_PUBCOMB.pep/\*  
 12: /cgn2\_6/prodata/1/pubpa/us09\_NEW\_PUB.pep/\*  
 13: /cgn2\_6/prodata/1/pubpa/us10A\_PUBCOMB.pep/\*  
 14: /cgn2\_6/prodata/1/pubpa/us10B\_PUBCOMB.pep/\*  
 15: /cgn2\_6/prodata/1/pubpa/us10C\_PUBCOMB..pep/\*  
 16: /cgn2\_6/prodata/1/pubpa/us10D\_PUBCOMB..pep/\*  
 17: /cgn2\_6/prodata/1/pubpa/us10 NEW\_PUB..pep/\*  
 18: /cgn2\_6/prodata/1/pubpa/us11\_NEW\_PUB..pep/\*  
 19: /cgn2\_6/prodata/1/pubpa/us11\_NEW\_PUB..pep/\*  
 20: /cgn2\_6/prodata/1/pubpa/us50\_NEW\_PUB..pep/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	210	100.0	38	14 US-10-092-750-46
2	58	27.6	1011	14 US-10-156-71-8999
3	57	27.1	184	14 US-10-080-110-608
4	57	27.1	184	16 US-10-080-110-608
5	57	27.1	184	17 US-10-168-316-608
6	54.5	26.0	102	16 US-10-437-963-127540
7	54	26.0	1755	16 US-10-437-963-186840
8	54	25.7	210	17 US-10-125-115-313042
9	54	25.7	275	14 US-10-106-658-51569
10	54	25.7	446	14 US-10-104-047-2206
11	54	25.7	484	15 US-10-382-122A-74677
12	54	25.7	517	14 US-10-094-749-2115
13	52	25.0	120	15 US-10-115-22000
14	53	25.2	378	16 US-10-437-963-159960
15	53	25.2	468	15 US-10-335-977-8588
16	53	25.2	522	15 US-10-333-977-8589
17	53	25.2	742	14 US-10-156-761-12153
18	53	25.2	1121	16 US-10-431-963-159359
19	52.5	25.0	88	15 US-10-424-599-176444
20	52.5	25.0	1439	16 US-10-437-963-138825
21	52	24.8	263	14 US-10-369-493-1.0395
22	52	24.8	1037	9 US-09-728-721-55
23	52	24.8	1037	9 US-09-728-721-55
24	52	24.8	1037	14 US-10-429-981-55
25	51.5	24.5	129	17 US-10-455-115-264527
26	51.5	24.5	261	14 US-10-192-381-44
27	51.5	24.5	382	14 US-10-360-493-14256
28	51	24.3	64	17 US-10-425-115-268841
29	51	24.3	135	9 US-09-783-436-12
30	51	24.3	486	15 US-10-282-122A-72523
31	51	24.3	486	15 US-10-148-90A-5
32	50.5	24.0	181	15 US-10-424-599-15802
33	50.5	24.0	190	14 US-10-050-882-57
34	50.5	24.0	311	15 US-10-232-122A-67409
35	50.5	24.0	336	14 US-10-050-170-25
36	50.5	24.0	336	16 US-10-080-170-5
37	50.5	24.0	336	17 US-10-468-356-25
38	50.5	24.0	509	14 US-10-369-493-11768
39	50.5	24.0	837	14 US-10-369-493-18586
40	50.5	24.0	1431	14 US-10-355-68A-38
41	50.5	24.0	1431	14 US-10-045-400C-5
42	50.5	24.0	1421	14 US-10-295-027-1235
43	50	23.8	93	16 US-10-431-963-15758
44	50	23.8	165	17 US-10-425-111-332283
45	50	23.8	637	15 US-10-425-114-57951

ALIGNMENTS

RESULT 1	US-10-092-750-46	;	Publication No. US20030032157A1
		;	GENERAL INFORMATION:
			APPLICANT: Hammoud, Philip W.
			APPLICANT: Alpin, Julia
			APPLICANT: Wright, Martin C.
			TITLE OF INVENTION: Peptptides Interactive with BCL-XL
			FILE REFERENCE: 50036/050002
			CURRENT APPLICATION NUMBER: US/10/092,750
			PRIOR APPLICATION NUMBER: 2002-01-07
			PRIOR FILING DATE: 2001-03-08
			NUMBER OF SEQ ID NOS: 253
			SOFTWARE: FastSEQ For Windows Version 4.0
			SEQ ID NO: 46
			LENGTH: 38
			TYPE: PRT
			ORGANISM: Homo sapiens
Qy	1 AAMLDLRGTECDLWINEMSLHKKIVQDVYGTPHPHHS	;	Sequence 8999, Application US/10156761
Dy	1 AAMLDLRGTECDLWINEMSLHKKIVQDVYGTPHPHHS	;	Sequence 8999, Application No. US03019018A1

SUMMARIES

Query Match	Best Local Similarity	Score 210;	DB 14;	Length 38;
Matches	Conservative	No. 2	No. 4	Gaps
1	100.0%	100.0%	0	0;

RESULTS

Result No.	Score	Query Match Length	DB ID	Description
1	210	100.0	38	14 US-10-092-750-46
2	58	27.6	1011	14 US-10-156-71-8999
3	57	27.1	184	14 US-10-080-110-608
4	57	27.1	184	16 US-10-080-110-608
5	57	27.1	184	17 US-10-168-316-608
6	54.5	26.0	102	16 US-10-437-963-127540
7	54	26.0	1755	16 US-10-437-963-186840
8	54	25.7	210	17 US-10-125-115-313042
9	54	25.7	275	14 US-10-106-658-51569
10	54	25.7	446	14 US-10-104-047-2206
11	54	25.7	484	15 US-10-382-122A-74677
12	54	25.7	517	14 US-10-094-749-2115
13	52	25.0	120	15 US-10-115-22000
14	53	25.2	378	16 US-10-437-963-159960
15	53	25.2	468	15 US-10-335-977-8588
16	53	25.2	522	15 US-10-333-977-8589
17	53	25.2	742	14 US-10-156-761-12153
18	53	25.2	1121	16 US-10-431-963-159359
19	52.5	25.0	88	15 US-10-424-599-176444
20	52.5	25.0	1439	16 US-10-437-963-138825
21	52	24.8	263	14 US-10-369-493-1.0395
22	52	24.8	1037	9 US-09-728-721-55
23	52	24.8	1037	9 US-09-728-721-55
24	52	24.8	1037	14 US-10-429-981-55
25	51.5	24.5	129	17 US-10-455-115-264527
26	51.5	24.5	261	14 US-10-192-381-44
27	51.5	24.5	382	14 US-10-360-493-14256
28	51	24.3	64	17 US-10-425-115-268841
29	51	24.3	135	9 US-09-783-436-12
30	51	24.3	486	15 US-10-282-122A-72523
31	51	24.3	486	15 US-10-148-90A-5
32	50.5	24.0	181	15 US-10-424-599-15802
33	50.5	24.0	190	14 US-10-050-882-57
34	50.5	24.0	311	15 US-10-232-122A-67409
35	50.5	24.0	336	14 US-10-050-170-25
36	50.5	24.0	336	16 US-10-080-170-5
37	50.5	24.0	336	17 US-10-468-356-25
38	50.5	24.0	509	14 US-10-369-493-11768
39	50.5	24.0	837	14 US-10-369-493-18586
40	50.5	24.0	1431	14 US-10-355-68A-38
41	50.5	24.0	1431	14 US-10-045-400C-5
42	50.5	24.0	1421	14 US-10-295-027-1235
43	50	23.8	93	16 US-10-431-963-15758
44	50	23.8	165	17 US-10-425-111-332283
45	50	23.8	637	15 US-10-425-114-57951

APPLICANT: OMURA, SATOSHI  
 / APPLICANT: IKEDA, HARUO  
 / APPLICANT: JUN  
 / APPLICANT: ISHIKAWA, JUN  
 / APPLICANT: HIRAKAWA, HIROSHI  
 / APPLICANT: SHIBA, TADAYOSHI  
 / APPLICANT: SAKAKI, YOSHITAKI  
 / APPLICANT: HATTORI, MASAHIRO  
 / TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 / FILE REFERENCE: 249-462  
 / CURRENT APPLICATION NUMBER: US/10/156,761  
 / CURRENT FILING DATE: 2002-05-29  
 / PRIOR APPLICATION NUMBER: JP 2001-204089  
 / PRIOR FILING DATE: 2001-05-30  
 / PRIOR FILING DATE: JP 2001-272697  
 / NUMBER OF SEQ ID NOS: 15109  
 / SEQ ID NO 8999  
 / LENGTH: 1011  
 / TYPE: PRT  
 / ORGANISM: Streptomyces avermitillii  
 us-10-156-761-8999

Query Match 27.6%; Score 59; DB 14; Length 1011;  
 Best Local Similarity 37.1%; Pred. No. 61;  
 Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

Qy 1 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 916 AHIIDQNSAAAM--TEVNAHRYTDKGSRHP 948

RESULT 3  
 US-10-080-170-608  
 / Sequence 608, Application US/10080170  
 / Publication No. US20030129601A1  
 / GENERAL INFORMATION:  
 / APPLICANT: COLE, S.T.  
 / TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR TREATMENT OF MYCOBACTERIOSIS  
 / FILE REFERENCE: 03495\_0218  
 / CURRENT APPLICATION NUMBER: US/10/080,170  
 / CURRENT FILING DATE: 2002-06-10  
 / PRIOR APPLICATION NUMBER: 60/270,123  
 / NUMBER OF SEQ ID NOS: 652  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 608  
 / LENGTH: 184  
 / TYPE: PRT  
 / ORGANISM: Mycobacterium tuberculosis  
 us-10-080-170-608

Query Match 27.1%; Score 57; DB 14; Length 184;  
 Best Local Similarity 35.3%; Pred. No. 13;  
 Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 2 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 122 AFILDRLMSQQVDADERRVALLRKTVGETWGLPSP 155

RESULT 4  
 US-10-080-170-608  
 / Sequence 608, Application US/10080170  
 / Publication No. US20040121322A9  
 / GENERAL INFORMATION:  
 / APPLICANT: COLE, S.T.  
 / TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR TREATMENT OF MYCOBACTERIOSIS  
 / FILE REFERENCE: 03495\_0218  
 / CURRENT APPLICATION NUMBER: US/10/080,170  
 / CURRENT FILING DATE: 2002-06-10  
 / NUMBER OF SEQ ID NOS: 652  
 / SEQ ID NO 127540

Query Match 27.1%; Score 57; DB 16; Length 184;  
 Best Local Similarity 35.3%; Pred. No. 13;  
 Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 2 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 122 AFILDRLMSQQVDADERRVALLRKTVGETWGLPSP 155

RESULT 5  
 US-10-468-356-608  
 / Sequence 608, Application US/10468356  
 / Publication No. US20040197856A1  
 / GENERAL INFORMATION:  
 / APPLICANT: COLE, STEWART  
 / TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR TREATMENT OF MYCOBACTERIOSIS  
 / FILE REFERENCE: 05394\_0019  
 / CURRENT APPLICATION NUMBER: US/10/468,356  
 / CURRENT FILING DATE: 2003-08-19  
 / PRIOR APPLICATION NUMBER: 10/080,170  
 / PRIOR FILING DATE: 2002-02-22  
 / PRIOR APPLICATION NUMBER: 60/270,123  
 / PRIOR FILING DATE: 2001-02-22  
 / NUMBER OF SEQ ID NOS: 655  
 / SOFTWARE: PatentIn Ver. 3.2  
 / SEQ ID NO 608  
 / LENGTH: 184  
 / TYPE: PRT  
 / ORGANISM: Mycobacterium tuberculosis  
 us-10-468-356-608

Query Match 27.1%; Score 57; DB 17; Length 184;  
 Best Local Similarity 35.3%; Pred. No. 13;  
 Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 2 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 122 AFILDRLMSQQVDADERRVALLRKTVGETWGLPSP 155

RESULT 6  
 US-10-437-963-127540  
 / Sequence 127540, Application US/10437963  
 / Publication No. US2004012334A1  
 / GENERAL INFORMATION:  
 / APPLICANT: La Rosa, Thomas J.  
 / APPLICANT: Kovacic, David R.  
 / APPLICANT: Zhou, Yihua  
 / APPLICANT: Wu, Wei  
 / APPLICANT: Boukharov, Andrey A.  
 / APPLICANT: Barbaauw, Brad  
 / APPLICANT: Li, Ping  
 / TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Title of Invention: Plants and Uses Thereof for Plant Improvement  
 / FILE REFERENCE: 38-21 (53221) B  
 / CURRENT APPLICATION NUMBER: US/10/437,963  
 / CURRENT FILING DATE: 2003-05-14  
 / NUMBER OF SEQ ID NOS: 204966  
 / SEQ ID NO 127540

Query Match 27.1%; Score 57; DB 14; Length 184;  
 Best Local Similarity 35.3%; Pred. No. 13;  
 Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 2 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
 Db 122 AFILDRLMSQQVDADERRVALLRKTVGETWGLPSP 155

LENGTH: 102  
 TYPE: PRT  
 ORGANISM: *Oryza sativa*  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1)..(102)  
 OTHER INFORMATION: unsure at all xaa locations

FEATURE: Clone ID: PAT\_MRT4530\_299882C.1.pep  
 US-10-437-963-127540

Query Match Score 54.5%; DB 16; Length 102;  
 Best Local Similarity 24.3%; Pred. No. 15;  
 Matches 9; Conservative 12; Mismatches 15; Indels 1; Gaps 1;

Qy 3 MLDRGTECDLWINEMS-LIHKIVQDDGTPHPHS 38  
 Db 14 LILDEASNGDPVYTHHQNQIESSEATLQXMHPHN 50

---

RESULT 7  
 US-10-437-963-188840  
 Sequence 188840, Application US/10437963  
 Publication No. US20040123343A1

GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO: 188840  
 LENGTH: 1755

TYPE: PRT  
 ORGANISM: *Oryza sativa*  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_85405C.1.pep  
 US-10-437-963-188840

Query Match Score 54.5%; DB 16; Length 1755;  
 Best Local Similarity 38.5%; Pred. No. 3.3e+08;  
 Matches 10; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

Qy 14 LWI-NEMSLHKKIVQDDVYGTTPHPHS 38  
 Db 1070 IWRGNNTALQHKLVNLHATPLGGHS 1095

---

RESULT 8  
 US-10-425-115-313042  
 Sequence 313042, Application US/10425115  
 Publication No. US20040214272A1

GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO: 313042  
 LENGTH: 210

; TYPE: PRT ;  
 ; ORGANISM: *Zea mays* ;  
 ; FEATURE: ;  
 ; OTHER INFORMATION: Clone ID: MRT4577\_48552C.1.pep ;  
 ; US-10-425-115-313042 ;  
 Query Match Score 54%; DB 17; Length 210;  
 Best Local Similarity 52.9%; Pred. No. 39;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Qy 22 LHKIVQDDVYGTTPHPHS 38  
 Db 1 MAYILDFFYTTCHPPHS 17

RESULT 9  
 US-10-106-698-5169  
 Sequence 5169, Application US/10106698  
 Publication No. US20030109630A1

GENERAL INFORMATION:  
 APPLICANT: Rubin, et al.  
 TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
 FILE REFERENCE: PA005P1  
 CURRENT APPLICATION NUMBER: US/10/106,698  
 CURRENT FILING DATE: 2002-03-27  
 PRIORITY NUMBER: PT/US00/26524  
 PRIOR FILING DATE: 2000-09-28  
 PRIORITY NUMBER: US 60/157,137  
 PRIOR FILING DATE: 1999-09-29  
 PRIORITY NUMBER: US 60/163,280  
 PRIOR FILING DATE: 1999-11-03  
 NUMBER OF SEQ ID NOS: 8564  
 SOFTWARE: PatentIn Ver. 3.0  
 SEQ ID NO 5169  
 LENGTH: 275

TYPE: PRT  
 ORGANISM: *Homo sapiens*  
 FEATURE:  
 NAME/KEY: MISC FEATURE  
 LOCATION: (154)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC FEATURE  
 NAME/KEY: MISC FEATURE  
 LOCATION: (181)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC FEATURE  
 LOCATION: (184)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: MISC FEATURE  
 LOCATION: (265)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 SEQ ID NO 5169

Query Match Score 54%; DB 14; Length 275;  
 Best Local Similarity 34.1%; Pred. No. 52;  
 Matches 14; Conservative 9; Mismatches 12; Indels 6; Gaps 3;

Qy 3 MLLDRGTECDL--WLNEM-SLLRKIVQDDVYGTTPHPHS 38  
 Db 3 LVLERSGDRLDEPDWLAQLRQLKEVKAGDI-GDBHPPTRS 42

RESULT 10  
 US-10-104-047-2206  
 Sequence 2206, Application US/10104047  
 Publication No. US2003236332A1

GENERAL INFORMATION:  
 APPLICANT: HELIX RESEARCH INSTITUTE  
 TITLE OF INVENTION: NO. US200323639A1 full length cDNA  
 FILE REFERENCE: HI-A0115  
 CURRENT APPLICATION NUMBER: US/10/104,047  
 CURRENT FILING DATE: 2002-03-25  
 PRIORITY NUMBER: ;  
 PRIOR FILING DATE: ;

NUMBER OF SEQ ID NOS: 4096  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 2206  
 LENGTH: 446  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-104-047-2206

RESULT 12  
 US-10-094-749-2115 ; Sequence 2115, Application US/10094749  
 ; Publication No. US003021971A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ISOGAI, TAKAO  
 ; APPLICANT: SUGIYAMA, TOMOYASU  
 ; APPLICANT: OTSUKI, TETSUJI  
 ; APPLICANT: WAKAMATSU, AI  
 ; APPLICANT: SATO, HIROYUKI  
 ; APPLICANT: ISHII, SHIZUKO  
 ; APPLICANT: YAMAMOTO, JUN-ICHI  
 ; APPLICANT: ISONO, YUTAKO  
 ; APPLICANT: HIO, YURI  
 ; APPLICANT: OTSUKA, KAORU  
 ; APPLICANT: NAGAI, KEIICHI  
 ; APPLICANT: IRIE, RYOTARO  
 ; APPLICANT: TAMECHIKI, TACHIRO  
 ; APPLICANT: SEKI, NAOKI  
 ; APPLICANT: YOSHIKAWA, TSUTOMU  
 ; APPLICANT: OTSUKI, MOTONUKI  
 ; APPLICANT: NAGAHARI, KENJI  
 ; APPLICANT: MASURO, YASUHIKO  
 TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
 FILE REFERENCE: 084335/0160  
 CURRENT APPLICATION NUMBER: US/10/094,749  
 CURRENT FILING DATE: 2002-03-12  
 PRIORITY NUMBER: 60/350,435  
 PRIORITY FILING DATE: 2002-01-24  
 PRIORITY APPLICATION NUMBER: JP 2001-328381  
 PRIORITY FILING DATE: 2001-09-14  
 NUMBER OF SEQ ID NOS: 3381  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 2115  
 LENGTH: 517  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-094-749-2115

Query Match Score 54; DB 14; Length 517;  
 Best Local Similarity 39.3%; Pred. No. 1e+02;  
 Matches 11; Conserv 1; Mismatches 16; Indels 0; Gaps 0;

Query 9 GTCGDIWINESSLHHKIVQDVGTPHP 36  
 Db 5 GARKQFWKRSNSKLPGSSIQHYGAQHP 32

RESULT 13  
 US-10-425-115-333859 ; Sequence 333859, Application US/10425115  
 ; Publication No. US004014272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rose, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; File Reference: 38-21153222; B  
 ; Current Application Number: US/10/425,115  
 ; Number of Seq ID Nos: 369326  
 ; SEQ ID NO: 333859  
 ; LENGTH: 303  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_67590C.1.pep  
 US-10-425-115-333859

Query Match Score 54; DB 15; Length 484;  
 Best Local Similarity 47.4%; Pred. No. 96;  
 Matches 9; Conserv 5; Mismatches 5; Indels 0; Gaps 0;

Query 13 DLWINENSLHHKIVQDVG 31  
 Software: PatentIn version 3.1  
 SEQ ID NO: 78614  
 LENGTH: 484  
 TYPE: PRT  
 ORGANISM: Streptococcus pyogenes  
 US-10-282-122A-74677

Query Match 25:2%; Score 53; DB 17; Length 303;  
 Best Local Similarity 37.9%; Pred. No. 79;  
 Matches 11; Conservative 6; Mismatches 8; Indels 4; Gaps 2;

Oy 12 CDLWNTNEMS-LIHKIVQD---VYGTPHPP 36  
 Db 5 CDDWLSTLGCLIFPSVLRPGLCYGTSDDPP 33

RESULT 14  
 US-10-437-963-159960  
 Sequence 159960, Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovacic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping  
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 201966  
 SEQ ID NO 159960  
 LENGTH: 378  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59285C.1.pep

Query Match 25:2%; Score 53; DB 16; Length 378;  
 Best Local Similarity 45.5%; Pred. No. 1e+02;  
 Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Oy 8 RGTECDLWINEMSLIHKIVQDV 29  
 Db 107 RSSFCCLWLNCFSNHKKIVQQI 128

RESULT 15  
 US-10-335-977-8588  
 Sequence 8588, Application US/10335977  
 Publication No. US20040052799A1  
 GENERAL INFORMATION:  
 APPLICANT: DOUGLAS SMITH et al.  
 TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES  
 RELATING TO HELICOBACTER PYLORI FOR  
 DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 10031  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 28 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD/ROM ISO9660  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: Windows NT 4.0  
 SOFTWARE: UNIX  
 CURRENT APPLICATION NUMBER: US/10/335,977  
 FILING DATE: 30-Dec-2002  
 PRIOR APPLICATION NUMBER: 08/993,002

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 9.32076 Seconds  
(without alignments)

392,268 Million cell updates/sec

Title: US-10-092-750-46

Perfect score: 210

Sequence: 1 AAMLDRRGTECDLWINEMSLIHKIVQDVYGTPHPPHS 38

Scoring table: BLCSUM62

Gapop 10.0 , Gapext 0.5

searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:

283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	57.5	27.4	308	2	D83858	hypothetical prote
2	57	27.4	574	2	T40213	hypothetical prote
3	57	27.1	184	2	G70956	hypothetical prote
4	57	27.1	1116	2	T20476	hypothetical prote
5	55	26.2	312	2	T20932	hypothetical prote
6	54	25.7	245	2	A99163	hypothetical prote
7	54	25.7	245	2	AH1124	hypothetical prote
8	53	25.2	522	2	B71978	hypothetical prote
9	52	24.8	131	2	AEF338	hypothetical prote
10	51.5	24.5	297	2	T0589	hypothetical prote
11	51	24.3	186	2	A81172	probable dCTP deam
12	51	24.3	259	2	A98302	oligopeptide trans
13	51	24.3	259	2	AG2981	hypothetical prote
14	50.5	24.0	181	2	G85330	interacts with sec
15	50.5	24.0	181	2	E91085	SyD protein [impor
16	50.5	24.0	181	2	A55944	syd protein - Esch
17	50.5	24.0	181	2	C86212	conserved hypothet
18	50.5	24.0	186	2	AE2649	alpha-glucosidase
19	50.5	24.0	561	2	D97431	probable alpha-glu
20	50.5	24.0	594	2	H84239	dimeethylsulfoxide
21	50	23.8	637	2	D61566	dCTP deaminase (EC
22	50	23.8	190	1	C84431	hypothetical prote
23	50	23.8	227	2	D71157	probable haloacid
24	50	23.8	336	2	D71157	hypothetical prote
25	50	23.8	300	2	T00374	probable membrane
26	50	23.8	361	2	H96334	hypothetical prote
27	50	23.8	459	2	AC0075	probable membrane
28	50	23.8	483	2	S75169	hypothetical prote
29	50	23.8	821	2	AE1855	hypothetical prote
					T16412	

## ALIGNMENTS

## RESULT 1

D83858

hypothetical protein BH1668 [imported] - Bacillus halodurans

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #text\_change 09-Jul-2004

C;Accession: D83858

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; His

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: D83858

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-308 &lt;STO&gt;

A;Cross-references: UNIPROT:Q9KCA8; GB:AP001512; GB:BA000004; PIDN:910174030; PIDN:BA805

A;Experimental source: Strain C-125

C;Genetics:

A;Gene: BH1668

Query Match 27.4%; Score 57.5; DB 2; Length 308;

Best Local Similarity 46.2%; Pred. No. 3.7; Matches 12; Conservative 5; Mismatches 4; Indels 5; Gaps 2;

Qy 13 DLWNEMSLIHKIVQDVYGTTPHPHS 38

Db 189 DLWTS-LCBHRI---YGTTHPNS 209

## RESULT 2

T40213

hypothetical protein SPBC31F10.10c - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: T40213

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.

Submitted to the EMBL Data Library, August 1997

A;Reference number: Z21913

A;Accession: T40213

A;Status: Preliminary

A;Molecule type: DNA

A;Map position: 2

C;Superfamily: Saccharomyces cerevisiae probable membrane protein YMR100w

A;Cross-references: UNIPROT:P87311; EMBL:297204; PIDN:CA100087.1; GSPDB:GN00067; SPDB:S

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references: 1-574 &lt;WOO&gt;

A;Experimental source: strain 972h-; cosmid c31F10

C;Genetics:

A;Gene: SPDB:SPBC31F10.10c

A;Cross-references:

Qy 3 MLUDRRGTCDCDLWINEMSLHAKIVQDVYGTPHPP 36  
 Db :||| :||| :||| :||| :||| :||| :||| :|||  
 21 VLYDRBALDCD--SEMSLNSLSHLVYLTSNSP 51

RESULT 3  
 G70956 hypothetical protein Rv3614c - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998  
 C;Accession: G70956  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feilwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A05000; MUID:98395987; PMID:9634230  
 A;Accession: G70956  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Residues: 1-184 <COL>  
 A;Cross-references: UNIPROT:O06269; GB:Z95557; PIDN:CA008952;  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv3614c

Query Match 27.1%; Score 57; DB 2; Length 184;  
 Best Local Similarity 35.3%; Pred. No. 2,4; Indels 0; Gaps 0;  
 Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Qy 2 AMILDRRSTECDLWINEMSLHAKIVQDVYGTPHPP 35  
 Db 122 AFIDRMSSQVDAEHRFLARRTVGETWGLDSP 155

RESULT 4  
 B70476 hypothetical protein ag\_2054 - Aquifex aeolicus  
 C;Species: Aquifex aeolicus  
 C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
 C;Accession: B70476  
 R;Deckert, G.; Warren, P.V.; Gaesteland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov. V.  
 Nature 392, 353-358, 1998  
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A;Reference number: A070300; MUID:98196666; PMID:9557320  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-116 <AGF>  
 A;Cross-references: UNIPROT:O67838; GB:AE000770; PIDN:AA007805.1; PIDN:g2984274;  
 A;Experimental source: strain VFS  
 C;Genetics:  
 A;Gene: ag\_2054

Query Match 27.1%; Score 57; DB 2; Length 116;  
 Best Local Similarity 41.9%; Pred. No. 18; Indels 12; Mismatches 6; Indels 0; Gaps 0;  
 Matches 13; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 6 DRRTTECDLWINEMSLHAKIVQDVYGTPHPP 36  
 Db 950 DKELETFDFYIATVSDGKIFLVLVHGTPBENP 980

RESULT 5  
 T20932 hypothetical protein F14H8.2 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T20932  
 R;Matthews, L.  
 submitted to the EMBL Data Library, October 1996

A;Reference number: 219348  
 A;Accession: T20932  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-312 <WIL>  
 A;Cross-references: UNIPROT:Q93482; EMBL:281061; PIDN:CAB02931.1; GSPPDB:GN00023; CEESP:F  
 A;Experimental source: clone F14H8  
 A;Gene: CESP\_F14H8.2  
 A;Map position: 5  
 A;Introns: 79/3; 136/3

Query Match 26.2%; Score 55; DB 2; Length 312;  
 Best Local Similarity 40.5%; Pred. No. 8,2; Mismatches 4; Indels 10; Gaps 2;  
 Matches 15; Conservative 4; Mismatches 10; Indels 8; Gaps 2;

Qy 2 AMILDRRSTECD-LWINEMSLHAKIVQDVYGTPHPP 36  
 Db 223 SVLIDMKPKRIDITTLWCGSEBVSL----DVPGP1PHRP 253

RESULT 6  
 A99163 hypothetical transport system ATP-binding protein [imported] - Agrobacterium tumefaciens  
 dipeptide transport system ATP-binding protein [imported] - Agrobacterium tumefaciens  
 C;Species: Agrobacterium tumefaciens  
 C;Accession: A99163  
 R;Goodney, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Mareelz, B.  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: A99163  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-245 <KUR>  
 A;Cross-references: UNIPROT:Q8U736; GB:AE007870; PIDN:AAK88827.1; PID:91515858; GSPPDB:  
 C;Genetics:  
 A;Gene: AGR\_L\_526  
 A;Map position: linear chromosome  
 Query Match 25.7%; Score 54; DB 2; Length 245;  
 Best Local Similarity 40.0%; Pred. No. 8,6; Mismatches 3; Indels 6; Gaps 1;  
 Matches 12; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 8 RGTTCDLWINEMSLHAKIVQDVYGTPHPP 37  
 Db 73 RGRE----ELKIAQMVFQDPYGSCLPRH 96

RESULT 7  
 AH3124 hypothetical protein dppF [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C;Species: Agrobacterium tumefaciens  
 C;Accession: AH3124  
 R;Hood, D.W.; Setubal, J.C.; Kau, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, G.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, F.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; E.W.  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AH3124  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-245 <KUR>  
 A;Cross-references: UNIPROT:Q8U736; GB:AB008689; PIDN:AAL45414.1; PID:917743114; GSPPDB:  
 C;Genetics:  
 A;Gene: dppF  
 A;Map position: linear chromosome

Query Match 25.7%; Score 54; DB 2; Length 245;  
Best Local Similarity 40.0%; Pred. No. 8.6;  
Matches 12; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Qy 8 RGTECDLWINEMSLHKLIVQDYYGTTPPPH 37  
Db 73 RGER-----ELKLAQNQFDQYGSILPRH 96

**RESULT 8**  
B71978 Hypothetical Protein jhp0054 - Helicobacter pylori (strain J99)  
A;Variety: strain J99  
C;Accession: B71978  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
P;Ailm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen number: A1800; PMID:9912055; PMID:9913682  
A;Accession: B71978  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-522 <ARN>  
A;Cross-references: UNIPROT:Q9ZN06; GB:AE001445; GB:AE001439; NID:94154559; PIDN:AAD0563  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp0054

Query Match 25.2%; Score 53; DB 2; Length 522;  
Best Local Similarity 54.2%; Pred. No. 27;  
Matches 13; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

Qy 2 AMILDRLRG--TECDLNEMSLH 23  
Db 338 AMLFDRGVETECDLKVDCKELLN 361

**RESULT 9**  
AR3338 Hypothetical protein EMEI0692 [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Accession: AR3338  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
P;DelVecchio, V.G.; Rapatral, V.; Redkar, R.J.; Parra, G.; Mujer, C.; Ilos, T.; Ivanova, I.; Mazur, M.; Gotsman, E.; Selkov, B.; Bizer, P.H.; Hagijs, S.; O'Callaghan, D.; Letessier, F.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A;Reference number: AD3252; PMID:11756688  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-131 <KDR>  
A;Cross-references: UNIPROT:QB8YHV4; GB:AE008917; PIDN:AAL51873.1; PID:gi17982624; GSPDB:G  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI0692  
A;Map position: I

Query Match 24.8%; Score 52; DB 2; Length 131;  
Best Local Similarity 34.8%; Pred. No. 8;  
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 13 DIWINEMSLHKLIVQDYYGTTPHP 35  
Db 61 ELWAKVIGLVLVIAAPHNGAPHP 83

**RESULT 10**  
T08589 Hypothetical protein L23H3.40 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: T08589  
R;Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Neves, H.W.; Mayer, K.F.X.; Schue, submitted to the Protein Sequence Database, May 1999  
A;Reference number: Z16098  
A;Accession: T08589  
A;Molecule type: DNA  
A;Residues: 1-297 <BEV>  
A;Cross-references: UNIPROT:Q9SSZ0; EMBL:AL050398; GSPDB:GN00062; ATSP:L23H3.40  
A;Experimental source: cultivar Columbia; BAC Clone L23H3  
C;Genetics:  
A;Gene: ATSP:L23H3.40  
A;Map position: 4  
A;Introns: 9/3; 44/3; 82/3; 142/3; 169/2; 196/2; 225/3; 252/3  
Query Match 24.5%; Score 51.5; DB 2; Length 297;  
Best Local Similarity 54.2%; Pred. No. 24;  
Matches 13; Conservative 2; Mismatches 6; Indels 3; Gaps 1;  
Qy 13 DLWNEMSLHKLIVQDYYGTTP 33  
Db 210 DLMKETSLNNSCSQFQDVIGTP 233

**RESULT 11**  
A81272 probable dCTP deaminase (EC 3.5.4.13) Cj1292 [imported] - Campylobacter jejuni (strain C;Species: Campylobacter jejuni  
C;Accession: A81272  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: A81272  
C;Keywords: C.Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilli, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Whitehead, S.; Barre, Nature 403, 665-668, 2000  
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals key A;Reference number: AB1250; PMID:10688204  
A;Accession: A81272  
A;Molecule type: DNA  
A;Residues: 1-186 <PAR2>  
A;Cross-references: UNIPROT:Q9PN07; GB:AL139078; GB:AL111168; NID:96968723; PIDN:CARB737  
A;Experimental source: serotype O2, strain NCTC 11168  
C;Genetics:  
A;Gene: dcd; Cj1292  
C;Superfamily: dCMP deaminase  
C;Keywords: hydrolase

Query Match 24.3%; Score 51; DB 2; Length 186;  
Best Local Similarity 42.1%; Pred. No. 16;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 9 GTBDLWNEMSLHKLIVQ 27  
Db 2 GLKADNWIRKMALEHKMIE 20

**RESULT 12**  
A98302 oligopeptide transport ATP-binding protein amf AGR L\_2744 [imported] - Agrobacterium t  
C;Species: Agrobacterium tumefaciens  
C;Accession: A98302  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C;Accession: A98302  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldmar, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Marketz, B.; Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A;Reference number: A97359; PMID:21608551; PMID:11743194  
A;Accession: A98302  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-259 <KUR>  
A;Cross-references: UNIPROT:Q8UAB8; GB:AE007870; PIDN:AAK89939.1; PID:gi15159894; GSPDB:  
C;Genetics:

A;Gene: AGR-L\_2744  
A;Map position: linear chromosome

Query Match 24.3%; Score 51; DB 2; Length 259;  
Best Local Similarity 33.3%; Pred. No. 24;  
Matches 12; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

Qy 3 MLLDRRGTECDLWINEMSLHLKIVODVYGTYPHHS 38  
Db 72 ILIDGRSSYDPLIERCMLQMVFQDPYGSUHQRS 107

RESULT 13  
A22981 hypothetical protein Atu3456 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AG2281  
C;Reference number: AB2577; MUID:21608550; PMID:11743133  
P;Wood, D.W.; Setubhai, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.  
erger, G.; Gillet, W.P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
i; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2321, 2001.  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Accession number: AB2577; MUID:21608550; PMID:11743133  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-259 <KUR>  
A;Cross-references: UNIPROT:Q8UAB8; GB:AE008689; PIDN:AAI4269.1; PID:gi17741855; GSPDB:Q  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu456  
A;Map position: linear chromosome

Query Match 24.3%; Score 51; DB 2; Length 259;  
Best Local Similarity 33.3%; Pred. No. 24;  
Matches 12; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

Qy 3 MLLDRRGTECDLWINEMSLHLKIVODVYGTYPHHS 38  
Db 72 ILIDGRSSYDPLIERCMLQMVFQDPYGSUHQRS 107

RESULT 14  
GB930 interacts with secY [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: G85910  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Iiiter, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimmlanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Accession number: AB5480; MUID:21074935; PMID:11206551  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-181 <STO>  
A;Cross-references: UNIPROT:Q8XET2; GB:AE005174; NID:gi12517269; PIDN:AAGG7907.1; GSPDB:Q  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: syd: Escherichia coli syd protein  
C;Superfamily: Escherichia coli syd protein

Query Match 24.0%; Score 50.5; DB 2; Length 181;  
Best Local Similarity 34.2%; Pred. No. 19;  
Matches 13; Conservative 4; Mismatches 14; Indels 7; Gaps 2;

Qy 1 AAMILDRRGTECDLWINEMSLHLK---IVQDVYGTYPH 35  
Db 6 AQALKDFTARYCDAWHEE---HKSWPLSEELYGVSP 39

Search completed: November 10, 2004, 14:52:16  
Job time : 10.3208 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

on: November 10, 2004, 13:38:57 ; Search time 50.6189 Seconds  
(without alignments)  
431.938 Million cell updates/sec

e: exact score: 210  
ence: 1 AAMLLDRGTECDLWINMSLHAKIVDQVYGTPHPPHS 38

hit table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

base : UniProt\_02:  
1: uniprot\_sprot:  
2: uniprot\_trembl:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

hit no.	Score	Query	Match	Length	DB	ID	Description
1	67	31.9	634	2	Q6C6T9	Yarrowia l	
2	61.5	25.3	167	2	Q93G79	rhodococcus	
3	58	27.6	537	2	Q6SGV8	uncultured	
4	58	27.6	537	2	AAS07882	Aas07882 uncultured	
5	58	27.6	1011	2	Q82N43	streptomyco	
6	57.5	27.4	308	2	Q9KCA5	Q9KCA5 streptomyco	
7	57.5	27.4	574	1	YB2A_SCHPO	bacillus h	
8	57	27.1	602	2	Q6E269	schizosacco	
9	57	27.1	184	2	Q7TW35	mycobacter	
10	57	27.1	302	2	Q81527	plasmidium	
11	57	27.1	302	2	Q81723	plasmidium	
12	57	27.1	430	2	Q961B9	Q961B9 drosophila	
13	57	27.1	1116	1	YK54_AQUAE	aquifex ae	
14	57	27.1	1568	2	Q8I332	plasmidium	
15	56.5	26.9	503	2	Q810Z0	Q810Z0 plasmidium	
16	55.5	26.4	431	1	ARGD_BIFLO	bifidobacter	
17	55	26.2	312	2	Q934B2	Q934B2 caenorhabd	
18	55	26.2	1251	2	Q9K4E9	streptomyco	
19	54.5	26.0	101	2	Q9CYU9	mus muscul	
20	54	26.0	756	2	Q7M762	Q7M762 mus muscul	
21	54.5	26.0	1317	2	Q9GQC3	Q9GQC3 brugia mal	
22	54.5	26.0	1608	2	Q9FW76	Q9FW76 oryza sati	
23	54	25.7	245	2	Q7CVF0	Q7CVF0 agrobacter	
24	54	25.7	245	2	Q8U736	Q8U736 agrobacter	
25	54	25.7	398	1	QXR1BUCA	QXR1BUCA bruchera a	
26	54	25.7	401	2	Q9RLB51	Q9RLB51 streptomyco	
27	54	25.7	484	2	Q99YK8	Q99YK8 streptocco	
28	54	25.7	486	2	Q8NZY7	Q8NZY7 streptocco	
29	54	25.7	517	1	CVO4_HUMAN	Q8wua7 homo sapi	
30	54	25.7	517	1	CAG3C298	CAG3C298 homo sapi	
31	54	25.7	517	1	C9DPC1	C9DPC1 brachypodium	

DT	01-DEC-2001 (TREMBLrel. 19; Created)						
DT	01-DEC-2001 (TREMBLrel. 19; Last sequence update)						
DT	01-OCT-2003 (TREMBLrel. 25; Last annotation update)						
DE	Benzene dioxygenase small subunit.						
GN	Name=body;						
OS	Rhodococcus sp. 19070.						
OC	Bacteria; Actinobacteriae; Actinomycetales;						
OC	Corynebacterineae; Nocardiidae; Rhodococcaceae;						
OX	NCBI_TaxID=161384;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=19070;						
RA	MEDLINE=21268842; PubMed=11375157;						
RA	Haddad S., Eby D.M., Neidle E.L.;						
RT	"Cloning and expression of the benzoate dioxygenase genes from Rhodococcus sp. strain 19070."						
RR	Rhodococcus sp. strain 19070.						
APB1	Environ. Microbiol. 67:2507-2514(2001).						
DR	EMBL: AF293441; RA58040.1; -.						
DR	GO: GO:0003824; EC: catalytic activity; IEA.						
DR	GO: GO:0016702; EC: oxidoreductase activity, acting on single d...; IEA.						
DR	GO: GO:0006725; EC: aromatic compound metabolism; IEA.						
DR	GO: GO:0006118; EC: electron transport; IEA.						
DR	InterPro; IPR00091; Ring_hydroxyl_B.						
DR	Pfam; PF00866; Ring_hydroxyl_B_1.						
KW	Dioxygenase.						
SQ	SEQUENCE 167 AA; 19861 MW; CADABCD83D692B91 CRC64;						
Query Match	29.3%; Score 61.5; DB 2; Length 167;						
Best Local Similarity	38.6%; Pred. No. 2, 4;						
Matches	17; Conservative 5; Mismatches 9; Indels 13; Gaps 2;						
QY	4 LLDRRGTCGDLWNEMNSLHK-----IVDGYGTPHP 35						
Db	102 ILBRRGTCGDLRNWFSLYRNTTDIYFGTNFVTLLVSG-PHP 144						
RESULT 3							
Q6SGV8	PRELIMINARY; PRT; 537 AA.						
AC	AAS07882; PRELIMINARY; PRT; 537 AA.						
DT	03-MAR-2004 (TREMBLrel. 27; Created)						
DT	05-JUL-2004 (TREMBLrel. 27; Last sequence update)						
DT	05-JUL-2004 (TREMBLrel. 27; Last annotation update)						
DT	05-JUL-2004 (TREMBLrel. 27; Last annotation update)						
DE	Hypothetical protein.						
GN	ORIGIN=EBAC080-L32B05.10;						
OS	uncultured bacterium 463.						
OC	Bacteria; environmental samples.						
OX	NCBI_TaxID=257394;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	Delong E.F.;						
RC	Submitted (DBC-2003) to the EMBL/GenBank/DBJ databases.						
[2]							
RP	SEQUENCE FROM N.A.						
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Delong E.F.;						
DR	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL: AY45861; AAS07882.1; -.						
DR	InterPro; IPR001005; Myb DNA binding.						
DR	PS00334; MYB_2; UNKNOWN_1.						
KW	Hypothetical Protein.						
SQ	SEQUENCE 537 AA; 58971 MW; 443FCC7B86231264 CRC64;						
Query Match	27.6%; Score 58; DB 2; Length 537;						
Best Local Similarity	45.8%; Pred. No. 29;						
Matches	11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;						
QY	12 CDLWNEMNSLHKYQDVYGTPHP 35						
Db	5 CSLWINHLD-HPIVKPSYPSPPEP 26						
RESULT 4							
AAS07882	PRELIMINARY; PRT; 537 AA.						
ID	AAS07882; PRELIMINARY; PRT; 537 AA.						
AC	AAS07882; PRELIMINARY; PRT; 537 AA.						
DT	03-MAR-2004 (TREMBLrel. 27; Created)						
DT	02-MAR-2004 (TREMBLrel. 27; Last sequence update)						
DT	02-MAR-2004 (TREMBLrel. 27; Last annotation update)						
DE	Hypothetical protein.						
GN	BBAC080-L32B05.10.						
OS	uncultured bacterium 463.						
OC	Bacteria; environmental samples.						
OX	NCBI_TaxID=257394;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	Delong E.F.;						
RC	Submitted (DBC-2003) to the EMBL/GenBank/DBJ databases.						
[2]							
RP	SEQUENCE FROM N.A.						
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Delong E.F.;						
DR	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL: AY45861; AAS07882.1; -.						
DR	InterPro; IPR001005; Myb DNA binding.						
DR	PS00334; MYB_2; UNKNOWN_1.						
KW	Hypothetical Protein.						
SQ	SEQUENCE 537 AA; 58971 MW; 443FCC7B86231264 CRC64;						
Query Match	27.6%; Score 58; DB 2; Length 537;						
Best Local Similarity	45.8%; Pred. No. 29;						
Matches	11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;						
QY	12 CDLWNEMNSLHKYQDVYGTPHP 35						
Db	5 CSLWINHLD-HPIVKPSYPSPPEP 26						

SQ	SEQUENCE	· 1011 AA;	111014 MW;	2ADBB2A2562166A9	CRC64;
	Query Match	27.6%;	Score 58;	DB 2;	Length 1011;
	Best Local Similarity	37.1%;	Pred. No 60;		
	Matches 13;	Conservative 8;	Mismatches 12;	Indels 2;	Gaps 1;
Qy	1 AAMILDRGTECDLWLNEMSLILHKIVQDVIGTGPFP	35			
Db	916 AHLUDQENSEAAM-~TEVNTAHLRYTDXGSRHP	948			
RESULT 6					
Q9KCA5	PRELIMINARY;	PRT;	308 AA.		
Q9KCA5	(TREMBrel. 15, Created)				
01-OCT-2000	(TREMBrel. 15, Last sequence update)				
01-OCT-2000	(TREMBrel. 24, Last annotation update)				
DT	BH1668 protein.				
DE	Name=BH1668				
GN	Bacillus halodurans.				
OS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
NCBI_TaxID=86665;					
[1]					
RA	Horikoshi K.;				
RA	"Complete Genome sequence of the alkaliphilic bacterium <i>Bacillus halodurans</i> and genomic sequence comparison with <i>Bacillus subtilis</i> .";				
RA	Nucleic Acids Res. 2:317-431 (2000).				
RA	Fujii F., Hirama C., Nakamura Y., Ogashawara N., Kuhara S.,				
RA	Takami H., Nakasone K., Takeki Y., Maeno G., Sasaki R., Masui N.,				
RA	Horiuchi K.;				
RA	"Complete Genome sequence of the alkaliphilic bacterium <i>Bacillus subtilis</i> .";				
RA	Nucleic Acids Res. 2:317-431 (2000).				
PIR:	D83858;				
InterPro:	IPR011099; Metalo_hydrolase				
SEQUENCE	308 AA;	35202 MW;	39DBB8710877DE77	CRC64;	
Query Match	27.4%;	Score 57.5;	DB 2;	Length 308;	
	Best Local Similarity	46.2%;	Pred. No. 18;		
	Matches 12;	Conservative 5;	Mismatches 4;	Indels 5;	Gaps 2;
Qy	13 DLWNEMSLILHKIVQDVIGTGPFPHS	38			
Db	189 DLNNTS-LCBAHRI----TGTHPNS	209			
RESULT 7					
YBA_SCHPO	STANDARD;	PRT;	574 AA.		
YBA_SCHPO	(Rel. 40, Created)				
AC	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	05-JUL-2004 (Rel. 44, Last annotation update)				
DT	Hypothetical protein C31F10_10 in chromosome II.				
GN	ORFNames=SPBC31F1.10c;				
CS	Schizosaccharomyces pombe (Fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OC	Schizosaccharomyces.				
NCBI_TaxID=4896;					
[1]					
SEQUENCE FROM N.A.					
STRAIN=722;					
MEDLINE=21848401;					
Wood V., Gwilliam R., Rajandream M.A., Lyne R., Stewart A., Seguros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feitshorn T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgeson G., Holroyd S., Hornsey T., Howard S., Huckle E.J., Hunt S., Jasels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Munagala K., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,					
RN	RP	SEQUENCE FROM N.A.			
[1]	NCBI_TaxID=1773;				
RA	STRAIN=H37RV;				
RA	Medline=9634230;				
RA	DOI=10.1038/31159;				
RA	Cole S.T., Broch R., Parkhill J., Garnier T., Churcher C.M., Harris D., Gordon S.V., Eiglmeier K., Gao S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsey T., Howard S., Huckle E.J., Hunt S., Jasels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Moule S., Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,				
RA	Murphy P.,				
RA	Bacterria; Actinobacteria; Actinomycetales; Actinomycetes;				
OC	Corynebacteriinae; Mycobacteriaceae; Mycobacterium.				
NCBI_TaxID=1773;					

RA	Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whithead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."; Nature 393:537-544 (1998).	ID Q81527; PRELIMINARY; PRT; 302 AA.
RT	SEQUENCE FROM N.A.	AC Q81527; PRELIMINARY; PRT; 302 AA.
RT	MEDLINE=22206494; PubMed=12218036;	DT 01-MAR-2003 (TREMBLrel. 23; Created)
RT	Pleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O., Peterson J.D., DeBont J.P., Dodson R.T., Gwin M.J., Hart D.H., Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Brmoleva M.D., Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Xhouri H.M., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.; "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains"; J. Bacteriol. 184:5419-5490 (2002).	DT 01-MAR-2003 (TREMBLrel. 23; Last sequence update)
RT	EMBL: BX842583; CAB08352.1; -; AE017171; AAK8075.1; -; PIR: GT0956; G70506.	DT 01-JUN-2003 (TREMBLrel. 24; Last annotation update)
RT	TIGR: MT3716; -; DR: TuberList; Rv3614C; -; KW: Complete proteome; Hypothetical protein.	DE Hydroxyethylthiazole kinase, Putative.
SQ	SEQUENCE 184 AA; 19834 MW; 632691F65DC5CA CRC64;	GN ORFNames=PP1920C; OS Plasmodium falciparum (isolate 3D7); OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. NCBI_TaxID=16329; RN [1]
RA	SEQUENCE FROM N.A.	RP SEQUENCE FROM N.A.
RA	Medline=12368864;	RX MEDLINE=12225705; PubMed=12368864;
RA	Gardner M.J., Hall N., Fung E., White O., Ber riman M., Hyman R.W., Nelson K.E., Bowman S., Paulsen I.T., James K., Carlton J.M., Pain A., Nelson C., Salzberg S.L., Craig A., Kyes S., Eisen J.A., Rutherford K., Shuh B., Peterson J., Angiuoli S., Chan M.S., Nene V., Shallom S.J., Stu B., Mather M.W., Vaidya A.B., Perez M., Allen J., Selangut J., Haft D., Mather M.W., Vaidya A.B., Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M., Barrell B.; "Genome sequence of the human malaria parasite Plasmodium falciparum."; Nature 419:510-511 (2002).	RX [2]
RA	RP SEQUENCE FROM N.A.	RX [2]
RA	Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M., Nakao B., Rowley D., Tamaki T., Wang F.; Davis R.W.; RA RAL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.	RX [2]
RA	EMBL: AE014850; RAN36470.1; -; HSPEP; P39593; IEQ.	RX [2]
RA	GO: 0004417; F: hydroxyethylthiazole kinase activity IEA.	DR GO: 0009328; P: thiamin biosynthesis; IEA.
RA	InterPro: IPR000417; Hyetylzyx_kinase.	DR GO: 0009328; P: thiamin biosynthesis; IEA.
RA	PFam: PF02110; HK. 1.	DR InterPro: IPR000417; Hyetylzyx_kinase.
RA	PRINTS: PRO0109; HYEHTZKNASE.	DR PFam: PF02110; HK. 1.
RA	KW Kinase.	RX [2]
RA	SQ SEQUENCE 302 AA; 33882 MW; C743CB0D6245EE CRC64;	RX [2]
RA	RP SEQUENCE 302 AA; 33882 MW; C743CB0D6245EE CRC64;	RX [2]
RA	Query Match Score 57; DB 2; Length 184; Best Local Similarity 35.3%; Pred. No. 12; Indels 0; Gaps 0; Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;	Query Match Score 57; DB 2; Length 302; Best Local Similarity 55.6%; Pred. No. 20; Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY	2 AMILDRRTEDCDWINENSLIHKIVQYVGYPHP 35	QY 20 SLIHKIVQDVYGTGPHPPH 37
DB	122 AFILDRLMSQQVDAEHHRVALLRKTVGETWGULPSP 155	DB 272 SLSHKIIDDIYVYSHNPH 289
RESULT 9	Q7TW35 PRELIMINARY; PRT; 184 AA.	RX [1]
ID	Q7TW35; PRELIMINARY; PRT; 184 AA.	RX [1]
AC	AC Q7TW35; PRELIMINARY; PRT; 184 AA.	RX [1]
DT	DT 01-OCT-2003 (TREMBLrel. 25; Created)	RX [1]
DT	DT 01-OCT-2003 (TREMBLrel. 25; Last sequence update)	RX [1]
DT	DT 01-OCT-2003 (TREMBLrel. 25; Last annotation update)	RX [1]
DB	DB Hypothetical protein Mb3644C.	RX [1]
GN	GN Orderedocustnames=Mb3644C;	RX [1]
OS	OS Mycobacterium bovis.	RX [1]
CC	CC Bacteria: Actinobacteridae: Actinomycetales:	RX [1]
OC	OC Corynebacterineae; Mycobacteriaceae;	RX [1]
OX	OX NCBI_TaxID=1765; RN [1]	RX [1]
RA	RA SEQUENCE FROM N.A.	RX [1]
RA	RA STRAIN=AF212/97; MEDLINE=22209107; PubMed=12788972;	RX [1]
RA	RA Garnier T., Bigliauer K., Camus J.-C., Medina N., Mansoor H., Simon S., Monsempe C., Simon S., Lacroix C., Grondjeit J., Atkin R., Mayes R., Keating L., Wheeler P.G., Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.; "The complete genome sequence of Mycobacterium bovis."; Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).	RX [1]
RA	RA EMBL: BX248346; CAD9530.1; -.	RX [1]
RA	RA Complete proteome; Hypothetical protein.	RX [1]
RA	RA SEQUENCE 184 AA; 19834 MW; 632691F65DC5CA CRC64;	RX [1]
RA	RA Query Match Score 57; DB 2; Length 184; Best Local Similarity 35.3%; Pred. No. 12; Indels 0; Gaps 0; Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;	RX [1]
QY	2 AMILDRRTEDCDWINENSLIHKIVQYVGYPHP 35	QY 20 SLIHKIVQDVYGTGPHPPH 37
DB	122 AFILDRLMSQQVDAEHHRVALLRKTVGETWGULPSP 155	DB 272 SLSHKIIDDIYVYSHNPH 289
RESULT 10	Q81527 PRELIMINARY; PRT; 302 AA.	RX [1]
RA	RA SEQUENCE FROM N.A.	RX [1]
RA	RA Tamai G., Pelle R., Muila F.J.; EMBL: AX168853; ANN6338.1; -.	RX [1]
RA	RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.	RX [1]
RA	RA HSSP; P39593; IEQ.	RX [1]
RA	RA GO: 0004417; F: hydroxyethylthiazole kinase activity IEA.	RX [1]
RA	RA InterPro: IPR000417; Hyetylzyx_kinase.	RX [1]
RA	RA PFam: PF02110; HK. 1.	RX [1]
RA	RA PRINTS: PRO109; HYEHTZKNASE.	RX [1]
RA	RA Kinase.	RX [1]
RA	RA SEQUENCE 302 AA; 33882 MW; 638E08852FC867F3 CRC64;	RX [1]
RA	RA [1]	RX [1]



Db	301 WIGEPSTIAHSILDQLPPTPH	320	RA	Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A., Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Leonard N., Line A., McLean J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabinowitz E., Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squires R., Stevens K., Taylor K., Tivey A., Unwin L., Woodward J., Sulston J.E., Craig A., Newbold C., Barrell B.G.; Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13. n.	
RESULT 13	YK54_AQUAE	STANDARD;	PRT;	1116 AA.	RA
AC	YK54_AQUAE				RA
DT	16-OCT-2001	(Rel. 40, Created)			RA
DT	16-OCT-2001	(Rel. 40, Last sequence update)			RA
DT	01-OCT-2004	(Rel. 45, Last annotation update)			RA
DE	Hypothetical protein AQ_2054.				RA
OS	OrderdLocusName=AQ_2054;				RL
OC	Aquifex_aeolicus;				DR
OC	Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.				DR
RN	[1]				EMBL; AL929356; CAD51803; 1; -.
RR	SEQUENCE FROM N. A.				EMBL; AL929356; CAD51803; 1; -.
RC	RESTRN				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
RX	MEDLINE=98196666; PubMed=9537320; DOI=10.1093/32331;				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
RA	Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.B., Overbeek R., Yooseph S., Nease M.A., Aujay M., Huber R., Feldman R.A., Short J.M., Olsen G.J., Swanson R.V., RT	"The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.",			SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
RA	Nature 392:353-358 (1998).				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isp-sib.ch).				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
CC	EMBL; AAC00770; AAC07805; 1; -.				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
DR	InterPro: IPR00770; B70476.				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
DR	PFam: PF04357; DUF490; 1.				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
KW	Complete proteome; Hypothetical protein; Transmembrane.				SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
FT	TRANSMEM	3			SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
FT	Potential.	20			SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
SQ	1116 AA;	127794 MW;			SEQUENCE 1568 AA; 18865 MW; F1D966566CB78DE7 CRC64;
Query Match	27.1%; Score 57;	DB 1;	Length 1116;		Query Match 26.9%; Score 56.5%; DB 2; Length 503;
Best Local Similarity	48.0%;	Pred. No. 44;			Best Local Similarity 39.4%; Pred. No. 44;
Matches	13;	Conservative			Matches 13; Conservative 4; Mismatches 15; Indels 1; Gaps 1;
DB	950 DRRGTECDLWINEMSLHLKIVQDVYCPHPP	36			DB
Qy	950 DKLTEFFYLNTVSDEYKIFLVLVHGTPENP	980			Qy
DR	Plasmodium falciparum (isolate 3D7), Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				DR
OX					OX
RN	[1]				RN
RP	SEQUENCE FROM N. A.				RP
RX	MEDLINE=22507960; PubMed=12620990;				RX
RA	Bortvin A., Egan K., Skaltscheva R., Akutsu H., Berry D.L., Yamagimachi R., Page D.C., Jaenisch R.; "Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei.", Development 130:1673-1680(2003).				RA
RA	DR	AF490340; AAO84498; 1; -.			RA
RT	MGI; MGI:24156377; Prame14.				RT
RT	SEQUENCE 503 AA; 58137 MW; ELABCOB005212B30 CRC64;				RT
RN	[1]				RN
RP	SEQUENCE FROM N. A.				RP
RX	MEDLINE=22507960; PubMed=12620990;				RX
RA	Bortvin A., Egan K., Skaltscheva R., Akutsu H., Berry D.L., Yamagimachi R., Page D.C., Jaenisch R.; "Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei.", Development 130:1673-1680(2003).				RA
RA	DR	AF490340; AAO84498; 1; -.			RA
RT	MGI; MGI:24156377; Prame14.				RT
RT	SEQUENCE 503 AA; 58137 MW; ELABCOB005212B30 CRC64;				RT
Query Match	26.9%; Score 56.5%; DB 2; Length 503;				Query Match 26.9%; Score 56.5%; DB 2; Length 503;
Best Local Similarity	39.4%; Pred. No. 44;				Best Local Similarity 39.4%; Pred. No. 44;
Matches	13;	Conservative			Matches 13; Conservative 4; Mismatches 15; Indels 1; Gaps 1;
DB	120 VLDIIRDAHQDFNDGAGSLHEVCSQDVFQKNQ	152			DB
Qy	4 LLDRRGTECDLWINEMSLHLKIV-QDVYCPHPP	35			Qy
DR	Search completed: November 10, 2004, 14:50:02				DR
Job Time :	52.6189 secs				Job Time : 52.6189 secs

GenCore Version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds  
(without alignments)  
181.178 Million cell updates/sec

Title: US-10-092-750-47  
Perfect score: 124  
Sequence: 1 PHQYKPIADLYRGRRESRPSAPR 22

Scoring table: BLOSUM62  
Gapext: 0.5  
Searched: 478119 seqs., 68318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing First 45 summaries

## Database : Issued Patents AA:\*

1: /cgn2\_6/ptcdatal1/iaa/5a\_COMB.pep:  
2: /cgn2\_6/ptcdatal1/iaa/5b\_COMB.pep:  
3: /cgn2\_6/ptcdatal1/iaa/6a\_COMB.pep:  
4: /cgn2\_6/ptcdatal1/iaa/6b\_COMB.pep:  
5: /cgn2\_6/ptcdatal1/iaa/pcrtus\_COMB.pep:  
6: /cgn2\_6/ptcdatal1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	40.3	293	4 US-09-540-236-2400	Sequence 2400, AP
2	48	38.7	485	4 US-09-540-19052	Sequence 19052, A
3	48	38.7	851	4 US-09-252-991A-19645	Sequence 19645, A
4	47	37.9	151	4 US-09-221-976-4254	Sequence 4254, AP
5	47	37.9	297	4 US-09-52-187-18764	Sequence 18764, A
6	47	37.9	502	4 US-10-14-002-548	Sequence 548, APP
7	46	37.1	186	4 US-09-252-991A-22673	Sequence 22673, A
8	46	37.1	285	4 US-09-252-991A-32954	Sequence 32954, A
9	46	37.1	372	4 US-09-252-991A-21599	Sequence 21599, A
10	46	37.1	458	4 US-09-22-991A-38585	Sequence 28585, A
11	46	37.1	724	4 US-09-56-737-22	Sequence 737, APPL
12	46	37.1	204	4 US-09-52-737-29	Sequence 29, APPL
13	45.5	36.7	204	4 US-09-513-999C-7997	Sequence 7997, APPL
14	45.5	36.7	204	4 US-09-513-999C-7998	Sequence 7998, APPL
15	45.5	36.7	219	3 US-08-224-747-4	Sequence 4, APPL
16	45.5	36.7	219	3 US-09-22-991A-8171	Sequence 4, APPL
17	45.5	36.7	219	3 US-09-247-373B-4	Sequence 4, APPL
18	45.5	36.7	326	4 US-09-252-991A-27158	Sequence 27158, A
19	45.5	36.7	1220	2 US-08-543-510B-36	Sequence 36, APPL
20	45.5	36.7	1220	4 US-09-536-728-32	Sequence 32, APPL
21	45	36.3	182	4 US-09-252-991A-21971	Sequence 27911, A
22	45	36.3	208	4 US-09-538-02-810	Sequence 810, APPL
23	45	36.3	724	4 US-09-56-737-27	Sequence 27, APPL
24	45	36.3	724	4 US-09-56-737-30	Sequence 30, APPL
25	44.5	35.9	904	4 US-09-52-991A-19257	Sequence 19257, A
26	44	35.5	79	4 US-09-189-039A-3889	Sequence 8889, APPL
27	44	35.5	190	4 US-09-252-991A-25320	Sequence 25320, A

## ALIGNMENTS

RESULT 1						
US-09-540-236-2400						
/ Sequence 2400, Application US/09540236						
/ Patent No. 5675910						
/ GENERAL INFORMATION:						
/ APPLICANT: Gary L. Breton et al.						
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT.						
/ FILE REFERENCE: 2709.2005-001						
/ CURRENT APPLICATION NUMBER: US/09/540-2336						
/ CURRENT FILING DATE: 2000-04-04						
/ NUMBER OF SEQ ID NOS: 3840						
/ SEQ ID NO: 2400						
/ LENGTH: 293						
/ TYPE: PRT						
/ ORGANISM: M. catarrhalis						
US-09-540-236-2400						

  

RESULT 2						
US-09-522-991A-19052						
/ Sequence 19052, Application US/09252991A						
/ Patent No. 6551195						
/ GENERAL INFORMATION:						
/ APPLICANT: Marc J. Rubenfield et al.						
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS						
/ FILE REFERENCE: 107196.136						
/ CURRENT APPLICATION NUMBER: US/09/252, 991A						
/ CURRENT FILING DATE: 1999-03-18						
/ LENGTH: 650						
/ TYPE: PRT						
/ ORGANISM: Aeruginosa for diagnostics and therapeutics						
US-09-252-991A-19052						

  

Query Match %						
Best Local Similarity						
Qy 2 WQYKPIADLYRGRRESRPS 19						
Db 217 WSMPVADRYLAQNSPA 234						
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;						

Best Local Similarity 66.7%; Pred. No. 46;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 8 ADLYRGRBESRPSAPR 22  
Db 436 ADLGRGPRSPARPR 450

RESULT 3  
US-09-252-991A-19645  
Sequence 19645, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196\_136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 19645  
LENGTH: 851  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19645

Best Local Similarity 40.9%; Pred. No. 61; Mismatches 4; Indels 7; Gaps 1;

Qy 2 WOYKPIADLYGRGR--SRPSAP 21  
Db 16 WPFRPLAQDFRNRSWPLRGAP 37

RESULT 4  
US-09-221-976-4254  
Sequence 4254, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET 054P2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent-pm  
SEQ ID NO 4254  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 6  
OTHER INFORMATION: Xaa = Phe,Ile,Leu,Val  
US-09-621-976-4254

Best Local Similarity 37.9%; Pred. No. 13; Mismatches 4; Indels 5; Gaps 0;

Qy 1 PWQYKPIADLYGRGR 17  
Db 84 PWESKNSTAVWRGRDSR 100

Query Match 37.9%; Score 47; DB 4; Length 151;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Query Match 37.9%; Score 47; DB 4; Length 502;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PWQYKPIADLYGRGR 17  
Db 297 PWESKNSTAVWRGRDSR 313

RESULT 7  
US-09-252-991A-22673  
; Sequence 22673, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; PRIOR APPLICATION NUMBER: US/09/252, 991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094, 190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO: 22673  
; LENGTH: 186  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22673

Query Match 37.1%; Score 46; DB 4; Length 186;  
Best Local Similarity 52.9%; Pred. No. 24; Mismatches 7; Indels 0; Gaps 0;

Qy 6 PIADLYGREGRESRSPSAPR 22  
Db 108 PAADAHGRGSSPQR 124

RESULT 8  
US-09-252-991A-32954  
; Sequence 32954, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; PRIOR APPLICATION NUMBER: US/09/252, 991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094, 190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO: 32954  
; LENGTH: 285  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32954

Query Match 37.1%; Score 46; DB 4; Length 285;  
Best Local Similarity 58.8%; Pred. No. 38; Mismatches 6; Indels 0; Gaps 0;

Qy 6 PIADLYGREGRESRSPSAPR 22  
Db 214 PAADPGPRRRPRAAPR 230

RESULT 9  
US-09-252-991A-27599  
; Sequence 27599, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252, 991A

CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094, 190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO: 27599  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27599

Query Match 37.1%; Score 46; DB 4; Length 372;  
Best Local Similarity 40.9%; Pred. No. 50; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PWQYKPIADLYGREGRESRSPSAPR 22  
Db 61 PCWRPNNGTGRAPTAAPR 82

RESULT 10  
US-09-252-991A-28585  
; Sequence 28585, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252, 991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094, 190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO: 28585  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28585

Query Match 37.1%; Score 46; DB 4; Length 458;  
Best Local Similarity 52.9%; Pred. No. 62; Mismatches 1; Indels 0; Gaps 0;

Qy 6 PIADLYGREGRESRSPSAPR 22  
Db 61 PURDPYRPRPPIBETPR 77

RESULT 11  
US-09-252-737-22  
; Sequence 22, Application US/09562737  
; Patent No. 6428967  
; GENERAL INFORMATION:  
; APPLICANT: Herz, Joachim  
; TITLE OF INVENTION: LDL Receptor Signalling Pathways  
; FILE REFERENCE: UTSW0108  
; CURRENT APPLICATION NUMBER: US/09/562, 737  
; CURRENT FILING DATE: 2000-05-01  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 22  
; LENGTH: 724  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Sequence  
US-09-562-737-22

Query Match 37.1%; Score 46; DB 4; Length 724;  
 Best Local Similarity 42.0%; Pred. No. 1e+02; Indels 0; Gaps 0;  
 Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
 Qy 1 PWQYKPIADLVRGRSRSRSPAP 21  
 Db 291 PRRYSPVAKDLRGEDIPREP 311

RESULT 14  
 US-09-513-999C-7998 ; Sequence 7998, Application US/09513999C  
 ; Patent No. 6783961 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; ATTORNEY: Duclert, A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961 ;  
 FILE REFERENCE: 59.US2.REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26  
 NUMBER OF SEQ ID NOS: 36681  
 SOFTWARE: Patent.pn  
 SEQ ID NO: 7998  
 LENGTH: 204  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-513-999C-7998

Query Match 36.7%; Score 45.5; DB 4; Length 204;  
 Best Local Similarity 40.9%; Pred. No. 31; Mismatches 7; Indels 1; Gaps 1;  
 Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;  
 Qy 2 WQYKPIADLYRG-RESRSPAPR 22  
 Db 28 WQYRQLSALHRAPRPTPDKAR 49

RESULT 15  
 US-08-924-747-4 ; Sequence 4, Application US/08924747  
 ; Patent No. 6063570 ; GENERAL INFORMATION:  
 ; APPLICANT: MCGRONIGLE, BRIAN  
 ; APPLICANT: O'KEEFE, DANIEL  
 ; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE  
 ; TITLE OF INVENTION: ENZYMES  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY  
 ; STREET: 1007 MARKET STREET  
 ; CITY: WILMINGTON  
 ; STATE: DELAWARE  
 ; COUNTRY: UNITED STATES OF AMERICA  
 ; ZIP: 19898  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: DISKETTE, 3.50 INCH  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
 SOFTWARE: MICROSOFT WORD VERSION 7.0A  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/924,747  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: FLOYD, LINDA AXAMETHY  
 REGISTRATION NUMBER: 33,692  
 REFERENCE/DOCKET NUMBER: CL-1108  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 302-892-8112  
 TELEFAX: 302-773-0164  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 219 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: protein  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:

Query Match 37.1%; Score 46; DB 4; Length 724;  
 Best Local Similarity 40.9%; Pred. No. 1e+02; Indels 0; Gaps 0;  
 Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 Qy 1 PWQYKPIADLVRGRSRSRSPAP 22  
 Db 291 PRRYSPVAKDLRGEDIPREP 312

RESULT 13  
 US-09-513-999C-7997 ; Sequence 7997, Application US/09513999C  
 ; Patent No. 6783961 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; ATTORNEY: Duclert, A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961 ;  
 FILE REFERENCE: 59.US2.REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26  
 NUMBER OF SEQ ID NOS: 36681  
 SOFTWARE: Patent.pn  
 SEQ ID NO: 7997  
 LENGTH: 204  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-513-999C-7997

Query Match 36.7%; Score 45.5; DB 4; Length 204;  
 Best Local Similarity 40.9%; Pred. No. 31; Mismatches 7; Indels 1; Gaps 1;  
 Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;  
 Qy 2 WQYKPIADLYRG-RESRSPAPR 22  
 Db 28 WQYRQLSALHRAPRPTPDKAR 49

Search completed: November 10, 2004, 14:55:35  
Job time : 9:10283 secs

*This Page Blank (USP10)*



APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants  
 FILE REFERENCE: 38-21(532)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 333161  
 LENGTH: 61  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_66957C.1.pep  
 US-10-425-115-333161

Query Match 2 WQYKPIADLYGRGRRSPSAPR 22  
 Best Local Similarity 44.4%; Score 55; DB 17; Length 61;  
 Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 SEQ ID NO 333161  
 LENGTH: 61  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_66957C.1.pep  
 US-10-425-115-333161

Query Match 5 WQAHPFKLFYGDGRPDPR 25  
 Best Local Similarity 42.9%; Score 44; DB 17; Length 61;  
 Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 SEQ ID NO 333161  
 LENGTH: 61  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_66957C.1.pep  
 US-10-425-115-333161

RESULT 3  
 US-10-156-761-13264  
 Sequence 13264, Application US/10156764  
 Publication No. US20030119018A1  
 GENERAL INFORMATION:  
 APPLICANT: OMURA, SATOSHI  
 APPLICANT: IKEDA, HARUO  
 APPLICANT: ISHIKAWA, JUN  
 APPLICANT: ISHIKAWA, HIROSHI  
 APPLICANT: SHIBA, TADAYOSHI  
 APPLICANT: SAKAKI, YOSHIOUKI  
 APPLICANT: HATTORI, MASAHIRA  
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 FILE REFERENCE: 249-262  
 CURRENT APPLICATION NUMBER: US/10/156,761  
 CURRENT FILING DATE: 2002-05-29  
 PRIOR APPLICATION NUMBER: JP 2001-204089  
 PRIOR FILING DATE: 2001-05-30  
 PRIOR APPLICATION NUMBER: JP 2001-272697  
 PRIOR FILING DATE: 2001-08-02  
 NUMBER OF SEQ ID NOS: 15109  
 SEQ ID NO 13264  
 LENGTH: 441  
 TYPE: PRT  
 ORGANISM: Streptomyces avermitilis  
 US-10-156-761-13264

Query Match 2 WQYKPIADLYGRGRRSPSAPR 22  
 Best Local Similarity 40.3%; Score 50; DB 14; Length 441;  
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 SEQ ID NO 13264  
 LENGTH: 441  
 TYPE: PRT  
 ORGANISM: Streptomyces avermitilis  
 US-10-156-761-13264

RESULT 4  
 US-10-117-700A-7  
 Sequence 7, Application US/10417700A  
 Publication No. US2004003358A1  
 GENERAL INFORMATION:  
 APPLICANT: ECOPIA BIOSCIENCES INC.  
 APPLICANT: ZAZOPoulos, Emmanuel  
 APPLICANT: STAFIA, Alfredo  
 APPLICANT: FARNET, Chris  
 TITLE OF INVENTION: Specialized dual condensation/epimerization domain in non-ribosom

APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Peptide synthetase systems  
 FILE REFERENCE: 3002-114US  
 CURRENT APPLICATION NUMBER: US/10/417,700A  
 CURRENT FILING DATE: 2003-04-17  
 NUMBER OF SEQ ID NOS: 139  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 7  
 LENGTH: 445  
 TYPE: PRT  
 ORGANISM: Actinoplanes sp.  
 US-10-417-700A-7

Query Match 6 PIADLYGRGRRSPSAPR 22  
 Best Local Similarity 58.8%; Score 50; DB 15; Length 445;  
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 Qy 6 PIADLYGRGRRSPSAPR 22  
 Db 410 PVAVYRPGRVSRPAAQR 426

RESULT 5  
 US-10-437-963-11116  
 Sequence 11116, Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 APPLICANT: Li, Ping  
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(532)1.B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO 11116  
 LENGTH: 147  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (1)-(147)  
 OTHER INFORMATION: unsure at all xaa locations  
 FEATUR  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_15123C.1.pep  
 US-10-437-963-11116

Query Match 7 PWQYKPIADLYGRGRRSPSAPR 22  
 Best Local Similarity 40.9%; Score 49; DB 16; Length 147;  
 Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
 Qy 7 PWQYKPIADLYGRGRRSPSAPR 22  
 Db 22 PWLWQPTSLHVLRPFPAPR 43

RESULT 6  
 US-10-369-491-15960  
 Sequence 15960, Application US/10369493  
 Publication No. US20030233675A1  
 GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hink, Gregory J.  
 APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.  
 APPLICANT: Chen, XianFeng  
 TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES



i APPLICANT: Mehraban, Fuad  
 i APPLICANT: Topper, James N.  
 i APPLICANT: Malyankar, Uriel  
 i APPLICANT: Wasserman, Scott  
 i APPLICANT: Edinger, Shlomit  
 i APPLICANT: Smithson, Glennita  
 i APPLICANT: Gunther, Erik  
 i APPLICANT: Komives, Laszlo  
 i TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 i CURRENT APPLICATION NUMBER: US/10/120,801  
 i CURRENT FILING DATE: 2002-04-11  
 i PRIOR APPLICATION NUMBER: 60/285748  
 i PRIOR FILING DATE: 2001-04-03  
 i PRIOR APPLICATION NUMBER: 60/286068  
 i PRIOR FILING DATE: 2001-04-24  
 i PRIOR APPLICATION NUMBER: 60/286292  
 i PRIOR FILING DATE: 2001-04-15  
 i PRIOR APPLICATION NUMBER: 60/288334  
 i PRIOR FILING DATE: 2001-05-03  
 i PRIOR APPLICATION NUMBER: 60/291241  
 i PRIOR FILING DATE: 2001-05-16  
 i PRIOR APPLICATION NUMBER: 60/322284  
 i PRIOR FILING DATE: 2001-09-14  
 i PRIOR APPLICATION NUMBER: 60/285609  
 i PRIOR FILING DATE: 2001-04-20  
 i NUMBER OF SEQ ID NOS: 155  
 i SOFTWARE: PatentIn Ver. 2.1  
 i SEQ ID NO: 71  
 i LENGTH: 901  
 i TYPE: PRT  
 i ORGANISM: human  
 i US-10-120-801-71  
 Query Match 39.5%; Score 49; DB 14; Length 901;  
 Best Local Similarity 42.9%; Pred. No. 2.8e+04;  
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 PWQYKPIADLYGRRESRSP 21  
 Db 873 PIKLSPTSPVYRSSSSGSSP 893  
 RESULT 11  
 US-10-120-801-70  
 i Sequence 70, Application US/10120801  
 i PUBLIC INFORMATION:  
 i APPLICANT: Pena, Carol  
 i APPLICANT: Guo, Xiaojia  
 i APPLICANT: Shimkess, Richard  
 i APPLICANT: Padigaru, Muralidhara  
 i APPLICANT: Kekuda, Ramesh  
 i APPLICANT: Sptyek, Kimberly  
 i APPLICANT: Mehraban, Fuad  
 i APPLICANT: Topper, James N.  
 i APPLICANT: Malyankar, Uriel  
 i APPLICANT: Wasserman, Scott  
 i APPLICANT: Edinger, Shlomit  
 i APPLICANT: Smithson, Glennita  
 i APPLICANT: Gunther, Erik  
 i APPLICANT: Komives, Laszlo  
 i TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 i CURRENT APPLICATION NUMBER: US/10/120,801  
 i CURRENT FILING DATE: 2002-04-11  
 i PRIOR APPLICATION NUMBER: 60/285748  
 i PRIOR FILING DATE: 2001-04-23  
 i PRIOR APPLICATION NUMBER: 60/286068  
 i PRIOR FILING DATE: 2001-04-24  
 i PRIOR APPLICATION NUMBER: 60/286292  
 i PRIOR FILING DATE: 2001-04-25  
 i PRIOR APPLICATION NUMBER: 60/288334  
 i CURRENT FILING DATE: 2002-04-11  
 i NUMBER OF SEQ ID NOS: 155  
 i SOFTWARE: PatentIn Ver. 2.1  
 i SEQ ID NO: 70  
 i LENGTH: 1163  
 i TYPE: PRT  
 i ORGANISM: human  
 i US-10-120-801-70  
 Query Match 39.5%; Score 49; DB 14; Length 1163;  
 Best Local Similarity 42.9%; Pred. No. 3.6e+02;  
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 Qy 1 PWQYKPIADLYGRRESRSP 21  
 Db 1135 PIKLSPTSPVYRSSSSGSSP 1155  
 RESULT 12  
 US-10-425-115-274406  
 i Sequence 274406, Application US/10425115  
 i Publication No. US20040214272A1  
 i GENERAL INFORMATION:  
 i APPLICANT: La Rosa, Thomas J.  
 i APPLICANT: Kovacic, David K.  
 i APPLICANT: Zhou, Yihua  
 i APPLICANT: Cao, Yongwei  
 i TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants  
 i FILE REFERENCE: 38-21(522)B  
 i CURRENT APPLICATION NUMBER: US/10/425,115  
 i CURRENT FILING DATE: 2003-04-28  
 i NUMBER OF SEQ ID NOS: 369326  
 i SEQ ID NO: 274406  
 i LENGTH: 84  
 i TYPE: PRT  
 i ORGANISM: Zea mays  
 i FEATURE:  
 i OTHER INFORMATION: Clone ID: MRT4577\_181849C.1.pep  
 i US-10-425-115-274406  
 Query Match 38.7%; Score 48; DB 17; Length 84;  
 Best Local Similarity 50.0%; Pred. No. 37;  
 Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
 Qy 1 PWQYKPIADLYGRRESRSP 22  
 Db 35 PHQQPAAPKTRGRSLSSPR 56  
 RESULT 13  
 US-10-322-797-3026  
 i Sequence 3026, Application US/10320797  
 i Publication No. US2004004955A1  
 i GENERAL INFORMATION:  
 i APPLICANT: Broskiin, Alexey M.  
 i APPLICANT: Zamudio, Carlos  
 i TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND  
 i FILE REFERENCE: 10182-021-999  
 i CURRENT APPLICATION NUMBER: US/10/320,797  
 i CURRENT FILING DATE: 2002-12-16  
 i PRIOR APPLICATION NUMBER: 60/3341,261  
 i PRIOR FILING DATE: 2001-12-17  
 i NUMBER OF SEQ ID NOS: 3361  
 i SOFTWARE: PatentIn version 3.1  
 i SEQ ID NO: 3026

Query Match 38.7%; Score 48; DB 15; Length 212;  
 Best Local Similarity 45.0%; Pred. No. 92;  
 Matches 9; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 2 WQKPIADLYGRGRRSPSAP 21  
 Db 28 WEYRQLAVTHRA--SRPSRP 45

RESULT 14  
 US-10-369-493-3763  
 ; Sequence 3763, Application US/10369493  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: PLANTS WITH IMPROVED PROPERTIES  
 ; CURRENT APPLICATION NUMBER: US/10/369493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO: 3763  
 ; LENGTH: 593  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1).:(593)  
 ; OTHER INFORMATION: unsure at all Xaa locations

Query Match 38.7%; Score 48; DB 14; Length 593;  
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;  
 Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 KPIADLVRGRGRRSPSAPR 22  
 Db 94 KPIFEFQQGNESPPPAR 111

RESULT 15  
 US-10-437-963-187550  
 ; Sequence 187550, Application US/10437963  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbausk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21 (53221) B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO: 187550  
 ; LENGTH: 125  
 ; TYPE: PRT



GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## Om protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds  
 (without alignments)  
 392.268 Million cell updates/sec

Title: US-10-092-750-47

Perfect score: 124

Sequence: 1 PWDKPTADLYRGRERSPSAPR 22

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79;\*

1: Pir1;\*

2: Pir2;\*

3: Pir3;\*

4: Pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	43.5	187	F72665	hypothetical prote
2	50.5	40.7	488	S74629	probable prote
3	50	40.3	217	A84806	probable retropro
4	50	40.3	282	A85076	hypothetical prote
5	48.5	39.1	204	S26380	ribosomal prote
6	48	38.7	432	T16141	hypothetical prote
7	47	37.9	1016	T31343	probline dehydrogen
8	46.5	37.5	323	T52297	hypothetical prote
9	46	37.1	213	E70633	transcription regu
10	46	37.1	290	AH2650	probable transcrip
11	46	37.1	290	G97432	probable phosphogl
12	46	37.1	632	C71327	hypothetical prote
13	46	37.1	679	HB4516	glucosidase BH0704
14	46	37.1	801	H82373	ribosomal protein
15	45.5	36.7	204	JG2369	conserved hypoth
16	45.5	36.7	306	H72417	SLN1 protein - Yea
17	45.5	36.7	1220	S48387	hypothetical prote
18	45	36.3	208	S59772	gene I47 protein -
19	45	36.3	247	S16869	alpha-soluble NSF
20	45	36.3	295	S58285	alpha-SNAP protein
21	45	36.3	295	G02238	beta-SNAP protein
22	45	36.3	295	S32367	choline dehydrogen
23	45	36.3	298	S22368	hypothetical prote
24	45	36.3	567	AC0143	probable glucosyl
25	45	36.3	574	T29615	ECF family sigma f
26	44.5	35.9	869	H83500	ecf sigma factor (
27	44	35.5	168	AB2759	hypothetical prote
28	44	35.5	177	C97540	hypothetical prote
29	44	35.5	312	A61183	hypothetical prote

## RESULT 1

F72665 hypothetical protein APE0749 - Aeropyrum pernix (strain KI)

Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C;Accession: F72665

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Take

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.;

DNA Res. 6, 83-10, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aerop

A;Reference number: A72450;

A;Accession: F72665

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-187 <KAN>

A;Cross references: UNIPROT:Q9Y221; DBJ:AP000060; NID:9510488; PMID:10382966

A;Experimental source: strain KI

C;Genetics:

A;Gene: APE0749

## ALIGNMENTS

Query	Match	%	Score	54;	DB 2;	Length
YKPADLYRGRERSPSAPR	22	5.9%	43.5	54;	DB 2;	Length 187;
YKPADLYRGRERSPSAPR	22	5.9%	43.5	54;	DB 2;	Length 187;

Best local Similarity 5.9%; Pred. NC. 0.91; Mismatches 1; Indels 0; Gaps 0;

Matches 11; Conservative

1; Mismatches

7; Indels

0; Gaps

0;

QY

YKPADLYRGRERSPSAPR

22

Db

YKVLGLLTYLGRERPLKLRP

53

RESULT 2

S74629

hypothetical protein S11686 - Synechocystis sp. (strain PCC 6803)

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

C;Accession: S74629

R;Kaneko, T.; Satoh, S.; Nakamura, A.; Asamizu, E.; Miyajima, N.

O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

B;Reference number: S74322; MUID:97061201; PMID:895231

A;Accession: S74629

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-488 <KAN>

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match % Score 50.5%; DB 2; Length 488;

Best Local Similarity 44.4%; Mismatches 5; Indels 1; Gaps 1;

A;Title: Structure of a gene in the dipteran Chironomus tentans encoding a yeast ribosomal protein L15 e - midge (Chironomus tentans)  
 A;Reference number: S26380; MUID:930652; PMID:1437565  
 A;Accession: S26380  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-204 <GAL>  
 A;Cross-references: UNIPROT:P30736; EMBL:X68332; NID:97116; PIDN:CAA48409.1; PID:97117  
 A;Superfamily: rat ribosomal protein L15  
 C;Keywords: protein biosynthesis; ribosome

RESULT 3  
 A84806 probable retroelement pol polyprotein [Imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Accession: AB4806  
 C;Date: 02-Feb-2001 #sequence\_change 09-Jul-2004  
 Qy 1 PWQYKPIADLYRGRERSP 18  
 Db 130 PWQYRPPS-IQLGRRPKP 146

Query Match 39.1%; Score 48.5%; DB 2; Length 204;  
 Best Local Similarity 45.5%; Pred. No. 6.8;  
 Matches 10; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Qy 2 WQYKPIADLYRG-RESEPSAPR 22  
 Db 28 WQYRQTKLHRPSPSPDKAR 49

RESULT 6  
 T16141 hypothetical protein F22F4.2 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Accession: T16141  
 C;Date: 20-Sep-1999 #sequence\_change 02-Sep-2000  
 A;Gene: At238520  
 R;Latreille, P.  
 A;Map position: 2  
 A;Description: The sequence of C. elegans cosmid F22F4.  
 A;Accession: T16141  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Accession: T16141  
 R;Latreille, P.  
 A;Map position: X  
 A;Description: The sequence of C. elegans cosmid F22F4.  
 A;Accession: T16141  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Introns: 8/2; 59/2; 386/3  
 C;Superfamily: Passover protein  
 C;Genetics:  
 A;Gene: CESP:F22F4.2  
 A;Map Position: X  
 A;Accession: T16141  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Introns: 8/2; 59/2; 386/3  
 C;Superfamily: Passover protein  
 C;Genetics:

RESULT 4  
 AB5076 hypothetical protein At4g07710 [Imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Accession: AB5076  
 C;Date: 16-Feb-2001 #sequence\_change 09-Jul-2004  
 Qy 1 PWQYKPIADLYRGRERSP 19  
 Db 178 PEDYKPIDQMGRDTPN 196

Query Match 40.3%; Score 50; DB 2; Length 217;  
 Best Local Similarity 47.4%; Pred. No. 4.3;  
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 WQYKPIADLYRG-RESEPSA--PR 22  
 Db 236 WGHTFADLYAGREWQDSGVFR 258

RESULT 7  
 T31343 proline dehydrogenase (EC 1.5.99.8) - Bradyrhizobium japonicum  
 C;Species: Bradyrhizobium japonicum  
 C;Accession: T31343  
 C;Date: 29-Oct-1999 #sequence\_change 09-Jul-2004  
 R;Straub, P.F.; Reynolds, P.H.; Althomsons, S.; Mert, V.; Zhu, Y.; Shearer, G.; Kohl, D.  
 Appl. Environ. Microbiol. 62, 221-229, 1996  
 A;Title: Isolation, DNA sequence analysis, and mutagenesis of a proline dehydrogenase gene  
 A;Accession number: 221008; MUID:9611420; PMID:8512700  
 A;Gene: T31343  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Introns: 1-1016 <STR>  
 A;Cross-references: UNIPROT:Q59206; EMBL:X90401; NID:g1212778; PIDN:CAA62038.1; PID:912  
 C;Genetics:  
 C;Superfamily: bifunctional proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehydrogenase  
 C;Keywords: oxidoreductase

RESULT 5  
 S26380 ribosomal protein L15 e - midge (Chironomus tentans)  
 N;Alternate names: ribosomal protein YL10 homolog  
 C;Species: Chironomus tentans  
 C;Accession: S26380; S25633  
 C;Date: 13-Jan-1995 #sequence\_change 13-Jan-1995  
 C;Cross-references: UNIPROT:Q59206; EMBL:X90401; NID:g1212778; PIDN:CAA62038.1; PID:912  
 R;Iacallì, J.; Wieslander, L.  
 Nucleic Acids Res. 20, 5473, 1992